جامعة مؤتة
عمادة الدراسات العليا

مقارنة صعوبة الفقرة وقدرة الفرد في ضوء كل من النظرية الكلاسيكية ونموذج راش

بسام احمد سالم المرافي

رسالة
مقدمة إلى
عمادة الدراسات العليا
استكمالا لمتطلبات الحصول على
درجة الماجستير في القياس والتقويم قسم علم النفس

جامعة مؤتة ، 2003م

جامعة مؤتة
عمادة الدراسات العليا

بسم الله الرحمن الرحيم

اجازة رسالة جامعية

تقرر اجازة الرسالة المقدمة من الطالب بسام احمد المرافي والموسومة بـ :

"مقارنة صعوبة الفقرة وقدرة الفرد في ضوء كل من النظرية الكلاسيكية ونموذج راش باستخدام فقرات الاختبار من متعدد للاختبار الوطني للصف التاسع الأساسي" استكمالا لمتطلبات الحصول على درجة الماجستير في القياس والتقويم.

القسم: علم النفس.

الأسم	التوقيع	التاريخ	
د. ساري سواقد		٢٠٠٣/٨/٢٦	مشرفاً
د. ماهر الدرابيع		٢٠٠٣/٨/٢٦	عضواً
د. عماد الزغول		٢٠٠٣/٨/٢٦	عضواً
د. عبداللة الصمادي		٢٠٠٣/٨/٢٦	عضواً

عميد الدراسات العليا

د. ذياب البداينة

الشكر والتقدير

أتوجه إلى الله سبحانه وتعالى بالحمد والثناء الذي أعانني على المواصلة والاستمرار حتى إنجاز هذا العمل .

ومن ثم أتقدم بالشكر والعرفان لكل من ساهم في الأشراف على هذه الرسالة منذ أن كانت فكرة حتى أصبحت واقعا . وأخص بالذكر الدكتور ساري سواقد الذي أشرف على هذه الرسالة وقدم لي النصح والإرشاد والتوجيه طيلة فترة عملي بها.

كما أقدم شكري وتقديري لكل من الدكتور عماد الزغول والدكتور عبد الله الصمادي والدكتور ماهر الدرابيع الذين كان لملاحظاتهم الدقيقة والقيمة أكبر الأثر في بلورة أفكار هذه الرسالة واثرائها.

وأقدم شكري وتقديري لكل من ساهم في تنفيذ إجراءات الرسالة من أصدقاء أعزاء وأخوة أوفياء .

وأخيرا أقدم شكري العميق وعظيم الأمتنان إلى الأهل الكرام : والديّ العزيزين صاحبي القلب الكبير ، وزوجتي الغالية والأبناء الأعزاء .

بسام احمد المرافي

جدول المحتويات

الموضوع **صفحة**

الفصل الثالث المنهجية والإجراءات

الفصل الرابع :عرض النتائج

الفصل الخامس :مناقشة النتائج والتوصيات

قائمة الجداول

قائمة الاشكال

قائمة الملاحق

الملخص

مقارنة صعوبة الفقرة وقدرة الفرد في ضوء كل من النظرية الكلاسيكية ونموذج راش

بسام أحمد سالم المرافي

جامعة مؤتة، 2003

هدفت هذه الدراسة الى المقارنة بين النظرية الكلاسيكية ونموذج راش في تقدير قدرة الفرد وقيم معاملات صعوبة الفقرات باستخدام فقرات الاختيار من متعدد للاختبار الوطني للصف التاسع في مبحث العلوم حيث بلغ عددها (20) فقرة.

وبالتحديد فقد حاولت الدراسة الأجابة عن الأسئلة الآتية:

1- هل توجد علاقة ارتباطية ذات دلالة احصائية($\alpha = 0.01$) بين تقديرات قدرة الفرد حسب النظرية الكلاسيكية ونموذج راش؟

2- هل توجد علاقة ارتباطية ذات دلالة احصائية ($\alpha = 0.01$) بين معاملات صعوبة الفقرات حسب النظرية الكلاسيكية ونموذج راش؟

3- هل قيم قدرات الافراد مستقرة عند استخدام عينات مختلفة من الفقرات ؟. وهل قيم معاملات الصعوبة للفقرات مستقرة عند استخدام عينات مختلفة من الافراد حسب النظرية الكلاسيكية ونموذج راش؟

تكون مجتمع الدراسة من طلبة الصف التاسع الاساسي في مديريات التربية والتعليم كافة والثقافة العسكرية والتعليم الخاص ووكالة الغوث في الاردن ، اما عينة الدراسية فقد تكونت من (5613) طالباً وطالبة موزعين على 164 مدرسة شملت مجموعة من مدارس المدن والريف.

تم تطبيق الاختبار الوطني على عينة الدراسة وتبين أن عدد الطلبة الذين خضعوا للاختبار قد بلغ (5520) طالباً وطالبة من أصل عينة الدراسة .

حللت إجابات أفراد الدراسة باستخدام برنامجي تحليل: " SPSS " للتحليل وفق النظرية الكلاسيكية ، وبرنامج " WinSteps " للتحليل وفق نموذج راش وذلك للحصول على الاحصائيات الخاصة بالفقرات والافراد .

أظهرت نتائج الدراسة إلى وجود اتفاق عالي بين النظرية الكلاسيكية ونموذج راش في توزيع قيم القدرة للطلبة المفحوصين وكذلك في تقدير معاملات صعوبة الفقرات.

وأظهرت النتائج أيضاً أن نموذج راش أظهر نوعاً من الاستقرار بين قيم قدرات الطلبة المفحوصين باستخدام اختبارين فرعيين أحدهما سهل والأخر صعب ، كما أظهر نموذج راش نوعاً من الاستقرار بين معاملات صعوبة الفقرات باستخدام عينتين إحداهما ذات أداء مرتفع والاخرى ذات أداء منخفض.، أي أن قدرة الفرد مستقلة عن الفقرات التي يختبر بها " Item-Free " ، وأن الفقرات مستقلة عن عينة الافراد المفحوصين "Person- Free "

أوصت الدراسة الباحثين ببناء اختبار تحصيلي وفق نموذج راش وعمل مقارنة لخصائص الفقرات وقدرات الافراد المفحوصين من خلال النظرية الكلاسيكية ونموذج راش .

كما اوصت الدراسة بالاستفادة من كل النظرية الكلاسيكية ونموذج راش في انتقاء الفقرات عند بناء المقاييس مستقبلاً.

Abstract

Comparison of item difficulty and person ability through

Classical test theory and Rasch model

Bassam Ahmad Al-Marafee

Mu'ta University, 2003

This study aimed at comparing between classical test theory and Rasch model in estimating person ability and item difficulty by using twenty multiple choice items of national 9th grade test in science.

Specifically , the study tried to answer the following questions :

Is there significant correlation between ability estimations according to classical test theory and Rasch model ?

Is there significant correlation between items difficulty estimations according to classical test theory and Rasch model?

Does persons ability be stable by using different samples of items? And does items difficulty be stable by using different samples of persons according to classical test theory and Rasch model?

The national test was applied to the sample which consisted of (5520) 9th grade students .

The responses were analyzed by (SPSS) program depending on classical test theory and (winsteps) program depending on Rasch model in order to obtain persons and items statistics.

Results showed that there is high agreement between classical test theory and Rasch model in person's abilities distribution. also in items difficulty distributions.

Also results showed that there is astable between students abilities in rasch model by using two subtest one is easy and the other is difficult.

And there is astable between items difficulties in rasch model by using two subsamples : high performance and low performance , this mean that Rasch model free the person ability from the effect of items and free the items difficulty from the effect of person ability.

The study recommended to construt achievement test according to rasch assumptions and comparing items and persons statistics by using classical test theory and rasch model.

It is also recomended to use both methodes: classicl test theory and rasch model in measurements construction in future.

الفصل الأول

خلفية الدراسة وأهميتها

مقدمة :

يسعى علماء القياس النفسي والتربوي إلى تحقيق الموضوعية في قياس الظاهرة السلوكية ، لأنه كلما كان القياس موضوعيا ، زادت الدقة في فهم الظاهرة موضوع القياس والتنبؤ بها ، وما يتبع ذلك من دقة في الضبط والتحكم التي تعتبر الغاية النهائية لأهداف العلم .

ويقصد بالموضوعية هنا عدم تأثر قياس الظاهرة السلوكية باختلاف الأداة المستخدمة في القياس ، وأن لا يتأثر أيضا بالافراد الذين يتم تقدير السمة لديهم بهذه الأداة . وأن يتم تدريج الأداة بوحدة قياس تتوافق مع تدريج مستويات الظاهرة السلوكية موضوع القياس كما يحدث في قياس الظواهر الفيزيائية ؛ فالتقدير الكمي لكتلة أحد الأجسام لايتغير بتغير الميزان المستخدم أو بتغير الأجسام التي توزن بهذا الميزان ، ولا يختلف في المعنى اذا عبرّنا عنه بوحدة الكيلوغرام أو اي وحدة أخرى من وحدات قياس الكتلة (كاظم ،1986) .

وعليه أشار كل من كروكر وألجينا (Crocker & Algina, 1986) إلى أهم الصعوبات التي تواجه القياس النفسي والتي قد تؤثر في موضوعيته وهي :

1- عدم وجود أساس نظري موحد ومعتمد في قياس سمة ما يمكن قبوله عالميا ، وذلك لأن القياس النفسي قياس غير مباشر يعتمد على السلوكات المرتبطة بالسمة نفسها . فقد يتفق عالمان على وجود سمة ما عند الفرد ولكنهما يختلفان في تحديد أنماط السلوكات التي تمثل تلك السمة ، مما يؤدي الى أختلاف في أجراءات القياس ، والتي قد تقود الى نتائج وتفسيرات مختلفة لمستوى امتلاك الفرد لتلك السمة .

2 – القياس النفسي عادة يعتمد على عينات محدودة من السلوكات لأنه من الصعوبة بمكان حصرجميع السلوكات المرتبطة بالسمة ، وقد لاتغطي هذه العينات مجال السمة موضوع الدراسة ، وبالتالي يتعذر وضع فقرات شاملة ومتنوعة لقياس السمة .

3- ما زال القياس النفسي يفتقر إلى وحدات قياس معرّفة تعريفا واضحا ومحددا . فمثلا لو أن مفحوصا ما لم يجب عن أي فقرة من فقرات اختبار يقيس القسمة الطويلـة فـي الرياضيات ، فهل هذا يعني انعدام تلك القدرة (السمة) لديه ؟ ولنفرض ان اختبارا ما ، تم تطبيقه على ثلاثة مفحوصين : أجاب الأول اجابة صـحيحةعن (5) فقـرات ، وأجـاب الثاني اجابة صحيحة عن (10) فقرات والثالـث أجـاب اجابة صحيحة عن (15) فقرة ، وكانت علامة كل فقرة من فقـرات الأختبار تساوي واحد ، فهل الفرق في الكفاية "Competence " بين المفحوصين الأول والثاني هي نفسها بين الثاني والثالث ؟ وهل المسافات بين المفحوصين الثلاثة متساوية على متصل السمة "Continuum " التي يقيسها الأختبار ؟.

4 – يصعب تعريف البناءات النفسية تعريفا إجرائيا على نحو مستقل ، دون تحديد علاقـتها مـع بناءات نفسية أخرى ذات صلة ، لذا أكد كل من لورد ونوفيك (Lord & Novick,1968) على أهمية تعريف البناء النفسي ضمن مستويين :

الأول : أن يعـرّف البناء النفسي بدلالة السلوك الملاحظ ، وهذا المستوى من التعريف يحدد اجراءات القياس .

الثانـي : ان يعـرّف البناء النفسي بدلالة علاقته المنطقية مع بناءات نفسية اخـرى ضـمن نظـام نظري " Theoretical system " ، وهذا المستوى من التعريف يساعدنا في تفسير نتائج القياس.

إن دراسـة الصعوبات السابقة التي تعترض القياس النفسي ووضع الحلول لها كانـت وما تزال من أولويات اهتمام علماء القياس الذين توجهوا للبحث عن منحى جديـد يخـتلف عـن الأتجاه المتداول حاليا الذي يستند الى النظرية الكلاسيكية في القياس "Classical test theory (CTT) " .

ومـــن الدوافـــع التـــي حفـــزت علماء القياس للبحث عن هذا المنحى الجديد والمغاير للنظرية الكلاسيكية هو وجود قصور"Shortcomings " في هذه النظرية والتـــي أشار إليها هامبلتون وسوامنثان (Hambleton & Swaminathan,1985) في النقاط الآتية :

1- عدم استقلال قيم إحصاءات فقرة الاختبار"Item statistics" مثل صعوبة الفقرة ومعـــامل تميـــيزها عن خصائص عينة الأفراد المفحوصين على فقرات ذلك الأختبار ، فإذا تغير أفراد العينة فإن هذه القيم سوف تتغير ، وكذلك احصاءات الفرد المفحوص " person statistics "غير مستقلة عن عينة فقرات الأختبار ، فـــاذا كـــان لدينا اختبار يقيس سمة معينة، فان علامة المفحوص على الأختبار تعـــتمد علـــى صعوبة فقراته ، بمعنى اذا كان الأختبار مكونا من فقرات سهلة نســـبيا ، فـــان العلامة الملاحظة ستكون عالية ، وهذا يعطي انطباعا بان قدرة المفحوص عالية ، واذا كان الأختبار مكونا من فقرات صعبة نسبيا ، فان علامة المفحـــوص ســـتكون منخفضـــة وبالتالـــي يعطي انطباعا بأن قدرة المفحوص منخفضة .

وبالمقابل اذا كان المفحوصون من ذوي القدرات العالية فان قيم معاملات صـــعوبة الفقرات ستكون عالية وهذا يعطي انطباعا أن فقرات الأختبار سهلة - على اعتبار ان معامل صعوبة فقرة ما يحدد بنسبة الذين اجابوا صحيحا عن تلك الفقـــرة - واذا كـــان المفحوصون من ذوي القدرات المتدنية فان قيم معاملات الصعوبة ستكون متدنية أي ؛ أن فقرات الأختبار تبدو صعبة .

2 – يـــتحدد موقـــع المفحوص على متصل القدرة أو السمة باستخدام فقرات اختبار معيـــن بالأوضاع والظروف التي تم فيها تطبيق الأختبار نفسه أو أختبار مواز له وعلى طبيعة الفقرات نفسها . فلو أن مفحوصين حصل كل واحد منهما على العلامـــة 50% علـــى اختباريـــن يقيســـان السمة نفسها ولكنهما مختلفان في الصعوبة ، فلا يمكن في هذه الحالة ان يكون المفحوصان متكافئين في القدرة ، وعليه فانه لايمكن التعامل مع المقارنة على أنها أخذت من أختبارات معيارية جيدة .

3 – مــن المفاهيم الأساسية للنظرية الكلاسيكية مفهوم الثبات والذي يعرّف بدلالة الاختــبارات المــتوازية أو اعادة الأختبار ، ولكن من الصعب تحقيق مفهوم "المتوازية" أو الشروط الواجب توفرها عند اعادة الأختبار من الناحية العملية ، لأن الأفــراد فــي التطبــيق الثاني للاختبار لا يمكن أن يكونوا هم أنفسهم في التطبيق الأول بسبب تأثرهم بعوامل التذكر أو النسيان والدافعية ...إلخ .

4 – لا تــزودنا هــذه النظرية بأي أساس للتنبؤ في كيفية أداء المفحوص على فقرة اختــبار مــا، بمعنى أنها لا تستطيع الإجابة عن السؤال : ما احتمال أن يجيب مفحوص ما أجابة صحيحة عن فقرة اختبار محددة ؟.

5 – تفــترض الــنظرية الكلاســيكية أن تبايــن أخطــاء القــياس هو نفسه لجميع المفحوصين .

ولتلافــي القصــور السابق وحل مشكلة الموضوعية في القياس ظهر الأتجاه الجديــد في القياس الذي يقوم على نماذج رياضية تعتمد على نظرية الاحتمالات ، والذي قام العالم لورد (Lord) في عام 1952 بوضع أسسه وافتراضاته وصاغه علــى شــكل نظرية سمّاها نظرية استجابة الفقرة " Item Response Theory" ، وفي عام (1960) قام عالم الرياضيات الدنماركي جورج راش (Rasch) بنشر أول نمــاذج هذه النظرية والذي يركّز على معلم واحد للفقرة وهو صعوبة الفقرة Item difficulty ، وسميّ هذا النموذج ب (نموذج راش) وطوّعه للتطبيق العملي العالم الأمريكــي بن رايت (Ben Wright) ، وبذلك بدأ القياس النفسي للظاهرة السلوكية يقــترب من القياس الفيزيائي الذي يتميز بعدم تأثر نتائج القياس بالأداة المستخدمة و بالعناصــر التــي اســتخدمت تلك الأداة طالما انها اداة مناسبة لقياس تلك الظاهرة . (Lord,1980)

وأصــبح بــناء المقاييس اما معتمدا على النظرية الكلاسيكية ومســتندا على افتراضــاتها ، أو معتمدا على نظرية استجابة الفقرة ومستندا على افتراضاتها، من هنا جاءت فكرة هذه الدراسة وهي مقارنة قدرة الفرد (المفحوص) وصعوبة الفقرة فــي ضوء النظرية الكلاسيكية ونموذج راش – أحد نماذج نظرية استجابة الفقرة –

مـــن خـــلال فقـــرات اختبار الاختيار من متعدد للأختبار الوطني في مبحث العلوم للصف التاسع الأساسي الذي تم تطبيقه في العام الدراسي 2000 / 2001 م .

مشكلة الدراسة :

تـــتفق كـــل من النظرية الكلاسيكية ونظرية استجابة الفقرة في أن الخصائص الســـيكومترية لفقـــرة الاختبار – والتي من أبرزها صعوبة الفقرة – هي التي تحدد العلاقة بين قدرة الفرد واحتمال إجابته عن تلك الفقرة .

نظـــريا إذا أجرينا مقارنة بين النظرية الكلاسيكية ونظرية استجابة الفقرة من خلال أحد نماذجها "نموذج راش" يتبين أنه بالرغم من اتفاقهما حول هذه الخصائص إلا أن هـــناك اخـــتلافا بينهما في كيفية معالجتهما لهذه الخصائص ، فمفهوم صعوبة الفقـــرة في إطار النظرية الكلاسيكية يدل على متوسط احتمالات الإجابة الصحيحة عن الفقرة لجميع أفراد عينة المفحوصين ، وتتغير قيمة صعوبة الفقرة بتغير العينة. وبالتالـــي فـــأن معرفتـــنا للقيمة التقديرية لصعوبة فقرة معينة بالنسبة لمجموعة من الأفـــراد لا يفيدنا كثيرا في معرفة درجة صعوبة الفقرة بالنسبة لأحد أفراد العينة ، ولا بالنســـبة لمجموعـــة أخرى من الأفراد،، وكذلك الحال بالنسبة لقدرة الفرد فهي لاتتحرر من عينة الفقرات التي تم أختباره بها.

أمـــا صـــعوبة الفقرة في نظرية استجابة الفقرة فهي النقطة على متصل السمة التـــي يكون احتمال اجابة الفرد – الذي قدرته تقع عند تلك النقطة – اجابة صحيحة يساوي (0.5) ، ويرجع ذلك الى أن نظرية أستجابة الفقرة تفترض وجود دالة مميزة لكـــل فقـــرة مـــن فقرات الأختبار تأخذ شكل المنحنى اللوغارتمي يســـمى منحنى خصـــائص الفقـــرة "ICC" . و هذا المنحنى يصف العـــلاقة بين أحتمال الأجابة الصحيحة عن الفقرة وموقع المفحوص على متصل السمة بحيث تتحرر قيم صعوبة الفقـــرات مـــن العيـــنة التـــي تـــم أختـــبارهـــا بهـــذه الفقـــرات "Sample-free" ، وكذلـــك تحـــرر قـــيم قدرات الأفراد من مجمـــوعة الفقرات المســـتخدمة Item-Free (Hambleton & Swamanthan, 1985) . وهذا يعني أن

الهـدف في هذه الحالة يكون موجها نحو الحصول على فقرات اختبار نستطيع من خلالها تقدير قدرات الأفراد المفحوصين .

وبالـرغم مـن القصـور في النظرية الكلاسيكية وخاصة المتعلق بعدم تحرر إحصـائيات فقـرة الاختـبار من عينة المفحوصين ، إلا أن أغلب المقاييس الحالية وخاصـة اختـبارات التحصـيل تبـنى علـى مفاهيم هذه النظرية ، بسبب بساطة افتراضاتها ونماذجها الرياضية .

لكـن كثيرا من الدراسات بينّت قدرة نظرية استجابة الفقرة على تلافي قصور الـنظرية الكلاسـيكية ممـا حـدا بالباحثين إلى توجيه دراساتهم نحو هذه النظرية باستخدام نماذجها المتعددة وخاصة نموذج راش الذي يمتاز ببساطته وتعدد تطبيقاته وسهولة استخدامه عن طريق برامج حاسوبية متعددة .

وأشـار فان (Fan,1998) إلى النقص الكبير في الدراسات التطبيقية التي تقارن بيـن الخصـائص السيكومترية لفقرات الاختبار وقدرة المفحوصين من وجهة نظر النظريتيـن معـا ، وأن هناك معرفة قليلة عن كيفية وإلى أي مدى تسلك النظريتان سلوكا مختلفا ؟ .

وعليه يمكن أن تتحدد مشكلة الدراسة في السؤال التالي:

هـل تخـتلف تقديـرات كل من قيم قدرات الأفراد وقيم معاملات صعوبة الفقرات باختلاف الإطار النظري لكل من النظرية الكلاسيكية ونظرية استجابة الفقرة "نموذج راش" ، وإلـى أي مدى ؟ وهل تتدرج هذه القيم- سواء صعوبة الفقرات أوقدرات الأفـراد- فـي الـنظرية الكلاسيكية كما تتدرج في نظرية استجابة الفقرة باستخدام نموذج راش ؟.

أهداف الدراسة

تسـلط هـذه الدراسة الضوء على أوجه الشبه والاختلاف بين النظريتين في رؤيـتهما لصعوبة الفقرة وقدرة الفرد من خلال فقرات اختبار تحصيلي ، وبالتالي فإن هدف الدراسة يمكن أن يتحدد في الأجابة عن الأسئلة الآتية :

1 – هـل توجـد علاقـة ارتباطية ذات دلالة احصائية ($\alpha = 0.01$) بين تقديرات قـدرات الأفـراد حسب كل من النظرية الكلاسيكية ونظرية استجـابة الفقرة "نموذج راش" ؟ .

2 – هـل توجـد علاقـة ارتباطية ذات دلالة احصائية($\alpha = 0.01$) بين معاملات صـعوبة الفقـرات حسب كل من النظرية الكلاسيكية ونظرية استجابة الفقرة "نموذج راش" ؟ .

3– هل قيم قدرات الأفراد مستقرة عند أستخدام عينات مختلفة من الفقرات؟ وهل قيم معـاملات الصـعوبة للفقرات مستقرة عند استخدام عينات مختلفة من الأفراد حسب كل من النظريتين؟ .

أهمية الدراسة :

تنـبع أهمـية هذه الدراسة من أن الاهتمام المتزايد من قبل علماء القياس في العلاقـة بين نظريتي القياس : النظرية الكلاسيكية ونظرية استجابة الفقرة يعود إلى أن دلالـة أي قياس يعتمد على خصائص نظرية القياس التي اشتق منها بحيث أولا: نتوقع من النظرية أن توضح كيف تم إنجاز القياس ، وثانيا : أن نكون قادرين على اختبار النظرية ومدى ارتباطها بالواقع .

وقـد حظيـت نظـرية أسـتجابة الفقـرة باهتمام علماء القياس المعاصرين والمؤسسـات العالمـية المعنـية ببـناء المقاييس لقدرتها في التغلب على كثير من المشكلات الناجمة عن أستخدام النظرية الكلاسيكية .

وبالـرغم من وضوح المعرفة النظرية للإطار النظري لكلا النظريتين إلا أن هـناك نقصا في المعرفة التطبيقية حول كيف وإلى أي مدى تسلك النظريتان سلوكا مختلفا .

ففـي هـذا الصدد أشار أكرمان (Ackerman,1999) بأنه إذا كان الهدف هو تحليل الخصائص الداخلية للاختبار فإن النظرية الكلاسيكية لا تكون كافية وفعّالة في هذه الحالة .

لذا جاءت هذه الدراسة للمقارنة بين النظرية الكلاسيكية ونظرية استجابة الفقرة مــن خلال أحد نماذجها الشائعة وهو نموذج راش الذي يفترض أن الفقرات تختلف فيما بينها في معامل الصعوبة فقط .

وتركز المقارنة على معلمين هما قدرة الفرد وصعوبة الفقرة والذي من خلال تفاعلهما تحدث استجابة الفرد ، وتستخدم هذه الدراسة بيانات حقيقية "Real Data" مــن نــتائج استجابات طلبة الصف التاسع الأساسي على فقرات الاختيار من متعدد للاختبار الوطني في مبحث العلوم .

الفصل الثاني

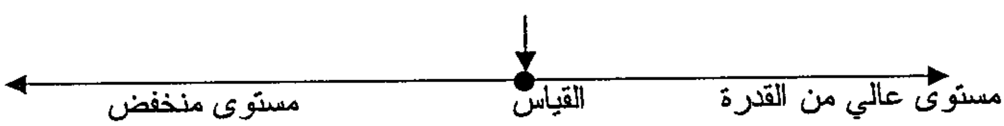
الاطار النظري والدراسات السابقة

الأطار النظري

تـتفق كـل من النظرية الكلاسيكية ونظرية استجابة الفقرة في افتراض وجود متصـل للسـمة "Continuum" ، وأنـه يمكـن تقدير احتمال إجابة فرد ما إجابة صـحيحة عـن فقرة اختبار إذا علمنا موقعه على هذا المتصل ، وأن هذا الاحتمال يزداد بازدياد مقدار السمة لديه (علام ، 1995) .

ويـرى عودة (1998) أنه يمكن استخدام السمة والقدرة كمصطلحين مترادفين مع أن القدرة تشير إلى خاصية عقلـية أو مهارية (نفس حركية) ، أما السمة فتشير إلـى خاصـية نفسـية أو انفعالـية . وعرّفـت أنستازي (Anastasi,1988) السمة بأنها : "مجموعة من السلوكات المترابطة التي تميل للحدوث معا" .

وبناء على اتفاق كل من النظريتين في افتراض وجود متصل للسمة أو القدرة فإنـه يمكـن تصوير نتيجة عملية القياس بتحديد نقطة تقع على خط مستقيم (متصل القـدرة) وذلـك من خلال اجابة الفرد (المفحوص) عن عينة من المثيرات (فقرات اختبار) التي تمثل السمة أو القدرة المرغوب قياسها كما تتضح في الشكل الآتي :



فـإذا اختـبرنا فرداً معينا باستخدام فقرات اختبار ما ، فإن نتيجة القياس هي تقديـر مكانـه على خط متصل القدرة ، والذي يحدده مقدار تفاعل قدرته مع فقرات ذلـك الاختبار ، على أساس أن الاختبار أداة قياس الغرض منه تحديد درجة امتلاك الفرد لسمة أو قدرة معينة .

وأشارت كاظم (1986) إلى أنه بالرغم من وجود العديد من العوامل المتداخلة التـــي قــد تؤثــر في استجابة الأفراد للفقرات غير عامل قدرة الفرد ، فإن ما يهمنا بـــالفعل تقدير مستوى قدرة الفرد ، وعلى هذا فمن المهم بذل الجهد وتنظيم الموقف الاختــباري لجعــل قدرة الفرد هي العامل الأساسي فقط الذي يتحكم في أدائه على الأختبار وتقليل آثار العوامل الأخرى المتداخلة ، وبالمثل بالنسبة لفقرات الاختبار ، فينبغـــي أيضـــا بذل أقصى الجهود لكي لا تكون هناك عوامل أخرى سوى صعوبة الفقـــرة تؤثر وتتحكم في التمييز بين الأفـــراد المختلفين في مستوى القدرة من خلال استجابتهم على هذه الفقرة .

إذن صعوبة الفقرة وقدرة الفرد هما العاملان الرئيسيان اللذان يجب أن يتحكما فـــي اســـتجابات الفـــرد علـــى فقرات اختبار ما كي نقترب من القياس الموضوعي للظاهـــرة الســـلوكية . وتخـــتلف نظرة كل من النظرية الكلاسيكية ونظرية استجابة الفقـــرة إلى قدرة الفرد وصعوبة الفقرة استنادا لافتراضات " Assumptions " كل منهما.

قدرة الفرد وصعوبة الفقرة في النظرية الكلاسيكية (CTT)

تقـــوم الـــنظرية الكلاســـيكية كمـــا يشـــير الـــيها هامبلـــتون وجونـــز (Hambleton & Jones , 1993) ، علـــى افتراضـــات نظـــرية بسيطة تتميز بعدم اســـتغراقها في نماذج رياضية معقدة مما يجعلها سهلة التطبيق في أوضاع اختبارية مختلفة ، وأهم هذه الأفتراضات هي مايلي :

1 – تـــتحدد درجة أو علامة المفحوص بمقدار ما يجيب إجابة صحيحة عن عدد مـــن الفقـــرات "الأسئلة" ، وهذه العلامة يرمز لها بالرمز (X) ، ويطلق عليها العلامة الملاحظة "Observed Score" التي يمكن أن تكون عرضة لوقوع الخطـــأ في تقديرها ، وهي تقدير لعلامة المفحوص الحقيقية (True Score) التي يرمز لها بالرمز (T) والتي تعبر عن قدرة المفحوص الفعلية .

وتتركب العلامة الملاحظة من جزأين هما:

العلامـــة الحقيقـــية *(T)* وهـــو مقدار يفترض بأنه ثابت ، والجزء الآخر هو مقدار الخطأ *(E)* ، حيث يتوزع هذا الخطأ توزيعا عشوائيا بوسط حسابي مقداره صفر.

والعلاقة بين العلامة الملاحظة و العلامة الحقيقية ومقدار الخطأ تمثل جميعا بالعلاقة التجميعية (Additive) الآتية (Lord & Novick, 1968) :

$$X = T + E$$

ويقصد بمقدارالخطأ الفرق بين العلامة الحقيقية (*T*) والعلامة الملاحظة (*X*) ، فالأخطـــاء هي عوامل تجعل العلامة الملاحظة أحيانا أعلى وأحيانا أقل من العلامة الحقيقـــية وعلـــى نحـــو عشـــوائي، لـــذا ســـمى هـــذا الخطـــأ بالخطـــأ العشوائي "Random Error" ، وقـــد يكـــون التأثـــير ثابتا أي في أتجاه واحد – بالزيادة أو النقصان – ويسمى الخطأ في مثل هذه الحالة بالخطأ المنتظم "Systematic Error" وعند الأشارة لخطأ القياس في النظرية الكلاسيكية يقصد به الخطأ العشوائي .

ولكـــي يكـــون مفهوم العلامة الحقيقية ومقدار الخطأ ذو معنى ، فأننا يجب أن نســـلّم بمـــثل هذا الأفتراض ، فاذا كانت العلامة الحقيقية للفرد تتغير باستمرار بين لحظـــة وأخـــرى فانـــه لايكون هناك أي معنى لقياسها ، لذلك من الضروري لهذه الـــنظرية أن تعتـــبر بـــأن الفـــرد يمتلك خصائص مستقرة ، ويمكن ان تكون هذه الخصائص قدرات او استعدادات أو ميول أوغير ذلك .

فالموضـــوع الأساســـي هنا هو أيا كانت الخاصية التي نهتم بقياسها يجب ان يمـــتلك الأفراد كمية مستقرة منها يعبّر عنها بالعلامة الحقيقية "T-SCORE "، وهذا يعنى أن النظرية تفترض ثبات السمة كي يمكن قياسها.

2- لا يوجـد ارتـباط بين قدرة الفرد والمعبّر عنها بالعلامة الحقيقية ((T ومقدار الخطأ العشوائي ((E ؛ أي أن :

$$\rho_{ET} = 0$$

حيث (ρ_{ET}) = معامل الارتباط بين العلامة الحقيقية والعلامة الخطأ.

وبسـبب وجود أخطاء في القياس بيّن لن (Linn, 1990) أن العلامة الملاحظة (X) يجب أن تفسّر بحذر ، وذلك كما أشار كل من ثورندايك وهيجين (1986) الى أن أي اختـبار مهمـا كان اسمه فإنه يقيس قدرات الشخص في الوقت الحاضر وما يستطيع الفرد القيام به "هنا" و"الآن" .

وقـد تـم توضيـح هذه القضيـة مـن قبـل علماء القياس بحيث لو أن مفحوصا تم اختباره في عشر فقرات من نوع الاختيار من متعدد ، ويعرف إجابة سبعا منها ولكنه أخطأ في اختيار اثنتين منها خطأ غير مقصود "Mismark " فإن علامته الملاحظة ستكون:

$$X = 7 - 2 = 5$$

ولو أن مفحوصا آخر أجاب عن أربع فقرات إجابة صحيحة وأجاب عن ثلاث فقـرات إجابة صحيحة عن طريق التخمين ، فإن علامته الملاحظة في هذه الحالة هي :

$$X = 4 + 3 = 7$$

ولنفرض أخيرا أن مفحوصا يعرف إجابة ثماني فقرات واخطأ في قراءةإحداها ولكنه خمّن تخمينا صحيحا عن فقرة أخرى فإن علامته في هذه الحالة هي :

$$X = 8 + 1 - 1 = 8$$

تبين هذه الأمثلة الثلاثة تأثير الخطأ العشوائي على العلامة الملاحظة – سواء في الاتجاه السلبي كما في الحالة الأولى أو في الاتجاه الإيجابي كما في الحالة الثانية أو كان التأثير صفرا كما في الحالة الثالثة (Crocker & Algina, 1986) .

لقـــد بيّن عودة (1998) أن الخطأ سواء كان موجبا أو سالبا فإن دقة القياس تـــتحدد بقيمة الخطأ وليس باتجاهه وأن أخطاء القياس مستقلة عن العلامة أو الدرجة الحقيقية ؛ بمعنى أننا لا نستطيع القول بأن من يأخذ علامة عالية يكون مقدار الخطأ في علامته أعلى أو أقل ممن يأخذ علامة منخفضة .

3- يشير الافتراض الثالث في النظرية الكلاسيكية إلى أنه إذا اختبرنا فرد ما بعدد غـــير محـــدد من المرات باستخدام الاختبار نفسه فإن الوسط الحسابي لعلامته الملاحظـــة هـــو تقديـــر لعلامته الحقيقية (T). والمشكلة هنا أنه من الصعب تطبـــيق الاختبار نفسه عدة مرات على المفحوص نفسه دون حدوث تداخل في آثار هذه التطبيقات على نتيجته ، لذا تبقى العلامة الحقيقية (T) علامة نظرية، أي لا نحصـــل عليها فعلا ولا يمكن قياسها قياسا تاما من العلامة الملاحظة ، وإنمـــا يمكن تقديرها ضمن مدى معين للعلامة الملاحظة . واعتمادا على ذلك فـــان تقدير قدرة الفرد من خلال النظرية الكلاسيكية يكون نتيجة تفاعل قدرته مع فقرات الأختبار في صورة أستجابة ملاحظة يعبر عنها بعلامة معينة.

أمـــا صـــعوبة الفقـــرة فـــي الـــنظرية الكلاسيكية فيذكر كروكر وألجينا (Crocker & Algina, 1986) على أنها "نسبة المفحوصين الذين أجابوا اجابة صـــحيحة عـــن الفقرة " ، واعتمادا على هذا التعريف ، فإن الفقرة التي أجاب عـــنها 85% مـــن المفحوصين فإن معامل الصعوبة (P) لها في هذه الحالة يســـاوي (0.85) ، بيـــنما الفقرة التي أجاب عنها 50% من المفحوصين فأن معـــامل صعوبتها يساوي (0.5) . وهذا يعني أنه كلما كانت الفقرة سهلة ، زاد معـــامل صعوبتها . ويبدو أن هذا يتعارض مع منطق مفهوم الصعوبة ، إلا أن هـــذا التعارض يزول إذا علمنا أن مصطلح "صعوبة الفقرة" هو مصطلح تقني "Term Technical " اتفق عليه علماء القياس .

وتعتقد ماتلوك (Matlock, 1997) أن صعوبة الفقرة (P) تتحدد من خلال أداء الأفراد (المفحوصين)على تلك الفقرة ، اذ من الصعوبة بمكان تحديد أي من سؤالين

مثلا هو الأكثر صعوبة بالنسبة للمفحوص من خلال قراءة هذين السؤالين فقط ما لم يتم الحصول على إجابة كل سؤال وتحديد قيمة معامل الصعوبة له .

تـتراوح قيمة (P) بين الصفر والواحد الصحيح ، فإذا كانت قيمة (P) تساوي صـفرا لفقـرة مـا ، فهذا يعني أن لا أحدا من المفحوصين أجاب عن هذه الفقرة ، بمعـنى أنه لا يوجد فروق فردية في العلامة على تلك الفقرة والمعنى نفسه لو كانت قيمة (P) تساوي واحد صحيح.

وإذا افترضنا أن عدد المفحوصين هو (N) وأن مجموع الدرجات التي حصل علـيها المفحوصـون على الفقرة (i) هو $\sum Xi$ ، فإن صعوبة هذه الفقرة يعبر عنها بالعلاقة :

$$Pi = \frac{\sum Xi}{N} * \frac{1}{Z}$$

حيث Z النهاية العظمى لعلامة الفقرة .

وعـندما تكـون الفقرة (i) مصححة ثنائيا (صفر أو واحد) فإن ΣXi سوف تكافـئ عـدد المفحوصين الذين أجابوا عن الفقرة اجابة صحيحة أي حصلوا على العلامة واحد على الفقرة (i) ، وعليه فإن معامل صعوبة الفقرة هو الوسط الحسابي لتلك الفقرة أي أن :

$$\mu i = \frac{\sum Xi}{N} = P_i$$

حيث (μi) الوسط الحسابي لعلامات المفحوصين على الفقرة i .

وإذا اعتبرنا أن معامل السهولة هو نسبة الطلبة الذين أجابوا إجابة خاطئة عن الفقـرة ، فـإن معـامل الصعوبة مضافا إليه معامل السهولة يساوي واحد صحيح وبالرموز :

$$P + q = 1$$

حيث (q) معامل السهولة وتساوي $(1-p)$

يلاحـــظ مـــن الجدول (1) أن معامل الصعوبة الأفضل يقترب من 0.50 كلما نقصت فرصة التخمين عن طريق زيادة عدد البدائل ، وهذه الطريقة ليست مضمونة النتائج اذ أن الإعداد الجيد للاختبار ودقة تنفيذه يقلل من ميل الطالب إلى التخمين .

لـــذا فـــإن معامل الصعوبة الأفضل في غياب التخمين هو (0.5) وبما أنه من الصـــعب أن تكون جميع الفقرات بهذا المستوى من الصعوبة ، فإن أي فقرة ضمن توزيـــع معـــاملات الصعوبة يتراوح بين (0.2) إلى (0.8) بمتوسط حسابي مقداره (0.5) يمكن أن تكون مقبولة (عودة ، 1986) .

أمـــا فـــي حالـــة الأختبارات محكية المرجع فأن تعريف معامل صعوبة الفقرة لايخـــتلف عن تعريفه في الأختبارات معيارية المرجع ، الا أن واضع اختبار محكي المـــرجع لايكون هدفه أختيار الفقرات لزيادة التباين ، وانما هدفه أنتقاء فقرات تقيس ناتج التعلم المقصود (Crocker & Algina, 1986).

وعلى الرغم من أن معامل الصعوبة يعد أبسط المقاييس الأحصائية استخداما ، الا أنـــه يعـــد فـــي واقع الأمر أكثرها صعوبة في التفسير . فطريقة حساب معامل الصـــعوبة لفقـــرة موضوعية يتطلب ايجاد عدد أفراد العينة الذين اجابوا عن الفقرة اجابـــة صـــحيحة ، الا أن تفسير القيمة الناتجة يكون في حدود هذه العينة والموقف التعليمـــي ، فاذا تغيرت العينة ربما تتغير هذه القيمة . كذلك اذا اختبر الأفراد عقب انـــتهاء التعليم مباشرة ، فان الفقرة تبدو سهلة بالنسبة لهم وبالتالي تتغير قيمة معامل صعوبة الفقرة وهذا التذبذب في قيم معاملات الصعوبة يجعل تفسيرها مربكا، ولعل اتجـــاه علمـــاء القياس الى المدخل الآخر الذي يعتمد على نظرية استجابة الفقرة قد ساعد في التغلب على هذه المشكلة (علام ، 1995) .

ويوضح ايبل (Ebel,1979) أنه من الصعب الحكم على فعالية الفقرة من خلال معـــامل صعوبتها فقط دون الاشارة الى تمييزها ، لأن لمعامل صعوبة الفقرة دلالة تميـــيز واضحة ؛ فالفقرة ذات التباين العالي ذات تمييز عالي ، وبالتالي فأن معامل الصعوبة الذي يعطي أعلى تباين هو المفضل . وهناك علاقة بين قيم معامل صعوبة

الفقرات ومعامل تمييزها ؛ فكلما ابتعدت قيمة معامل الصعوبة عن(0.5) ، فان قيمة معامل التمييز تقل ، لذا يكون تمييز الفقرات السهلة والفقرات الصعبة ضئيلا .

قدرة الفرد وصعوبة الفقرة في نظرية استجابة الفقرة (IRT)

تقـــوم نظرية استجابة الفقرة على عدة افتراضات أساسية أشار اليها هامبلتون وآخرون (Hambleton.,Swaminathan,& Rogers, 1991) على النحو الآتي :

الافتراض الأول :

أحادية البعد للسمة الكامنة Unidimensionality Latent Trait

ويعنـــي هـــذا الافتراض أن احتمال إجابة الفرد عن الفقرة يتحدد بدلالة السمة الكامنة لديه أو ما يسمى قدرة الفرد "Person Ability" وأن هناك قدرة واحدة فقط تفســـر أداء الفرد على الفقرة . وبالرغم من وجود عوامل اخرى بجانب قدرة الفرد كـــالعوامل الشخصية والعقلية وعوامل تتعلق بظروف تطبيق الأختبار والتي قد تؤثر في الأداء على الأختبار ، الا أن هذا الأفتراض يتطلب وجود عامل أساسي يؤثر في الأداء على الأختبار وهو مايشار اليه بالقدرة التي تقيسها فقرات الأختبار.

وقـــد اتفـــق كـــل مـــن لومســـدن وهامبلـــتون وتـــروب ولـــورد ونوفـــيك Lumsden,1961; Hambleton and Traub,1971;Lord and Novick,1968 المشار اليهم في (Hambleton and Swaminthan, 1985) على أن التحليل العاملي يستخدم في فحص أحادية البعد لمجموعة من الفقرات .

الافتراض الثاني :

الاستقلال المحلي Local Independence

يشـــير هـــذا الافتراض الى أن استجابات المفحوصين على فقرات مختلفة من اختبار ما مستقلة إحصائيا ، أي أن أداء المفحوص على فقرة معينة يجب أن لا يؤثر فـــي استجابته على فقرة أخرى وهذا يقتضي أن المحتوى لفقرة ما يجب أن لا يزوّد

المفحـــوص بإجابـــات عن فقرات أخرى ، على أفتراض أن كل فقرة مســـتقلة في بنائها ، ويتحقق أفتراض الاستقلال المحلي بتحقق افتراض احادية البعد للسمة.

الافتراض الثالث

منحنى خصائص الفقرة Item Characteristic Curve (ICC)

يشير هذا الافتراض الى أن احتمال حصول الفرد على الأجابة الصحيحة على فقـــرة اختبارية تكون دالة (Function) لكل من القدرة – التي يفترض أن الاختبار يقيسها لدى الفرد – وخصائص الفقرة التي يحاول الإجابة عنها وبالتالي فإن نظرية اســـتجابة الفقرة تهدف للتوصل إلى قيم تقديرية أحداها تتعلق بالفرد والآخرى تتعلق بالفقـــرة التي يختبر بها ، ثم تستخدم هذه القيم في تقدير احتمال الاستجابة الصحيحة لكــــل فقـــــرة مــــن فقـــرات الاختبار من خلال منحنى يسمى منحنى خصائص الفـــقـــرة "Item Characteristic Curve "ICC) أو اقـــتران اســـتجابة الفقرة (Item Response Function)، الذي يحدد العلاقة مابين قدرة الفرد المفحوص(θ) واحتمال حصوله على الأجابة الصحيحة عن فقرة ما (p) ، هذه العلاقة تتخذ شكل المنحنى اللوغارتمي على شكل الحرف (S) .

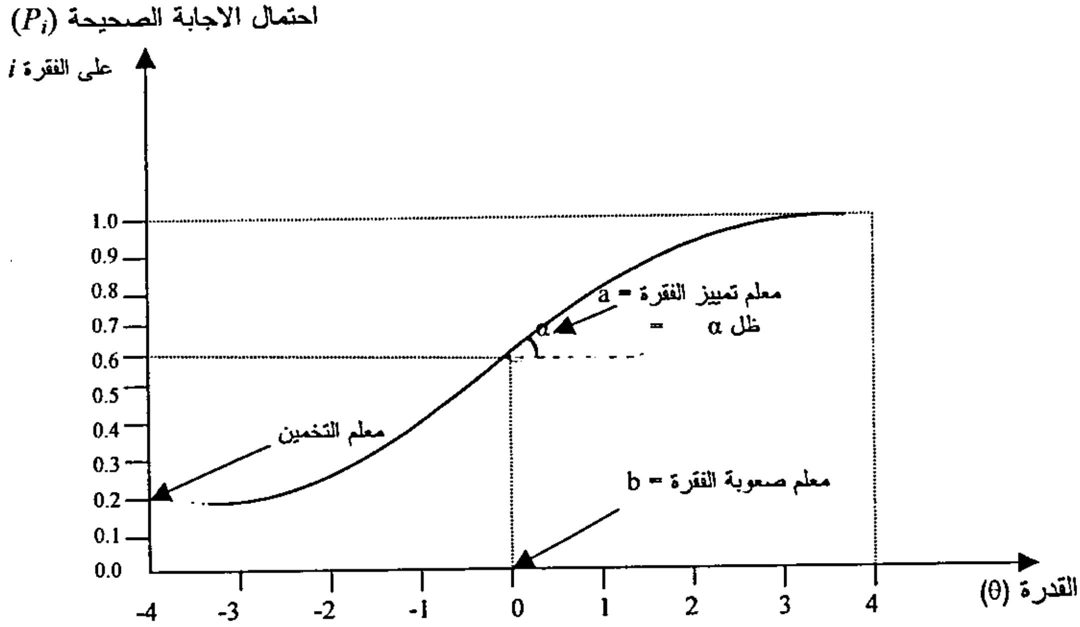
ويوصف المنحنى بدلالة ثلاثة أنواع من المعالم "Parameters" هي :

معلم الصعوبة (b) ، ومعلم التمييز (a) ، ومعلم التخمين (c) للفقرة .

واســـتنادا الــــى الدراســـات حول نظرية استجابة الفقرة ظهرت عدة نماذج لها تختلف فيما بينها في عدد المعلمات ، وأكثر النماذج شهرة في هذه النظرية هي :

1- النموذج اللوغارتمي ثلاثي المعالمThree- Parameter Logistic Model

ويسمى ايضا نموذج بيرنباوم (Birnbaum Model)، وهذا النموذج يأخذ بعين الأعتبار ثلاثة معالم للأداء على الفقرة هي: الصعوبة و التمييز والتخمين كما هو موضح في الشكل (1):



الشكل رقم (1)

منحنى خصائص الفقرة (i) ذو المعالم الثلاث

وبسبب وجود تخمين ، فان المنحنى في شكل (1) يحدث له أزاحة للأعلى بحيث تكون النقطة على محور(P) أعلى من (0.5) ، أما مقدار التمييز للفقرة a) فيقدر بميل المنحنى عند هذه النقطة والذي تمثله المعادلة التالية:

$$[a = \tan\alpha = slope]$$

ومعادلة هذا النموذج هي :

$$Pi(\theta) = C_i + (1 - C_i)\frac{e^{a(\theta - bi)}}{1 + e^{a(\theta - bi)}}$$

حيث :

$Pi\ (\theta)$ = احتمال أن يجيب الفرد (المفحوص) اجابة صحيحة عن الفقرة (i).

θ = مقدار القدرة لدى الفرد كما يقيسها الأختبار .

C_i = احتمال ان يجيب الفرد اجابة صحيحة على الفقرة(i) من خلال التخمين .

a_i = تمييز الفقرة (i) .

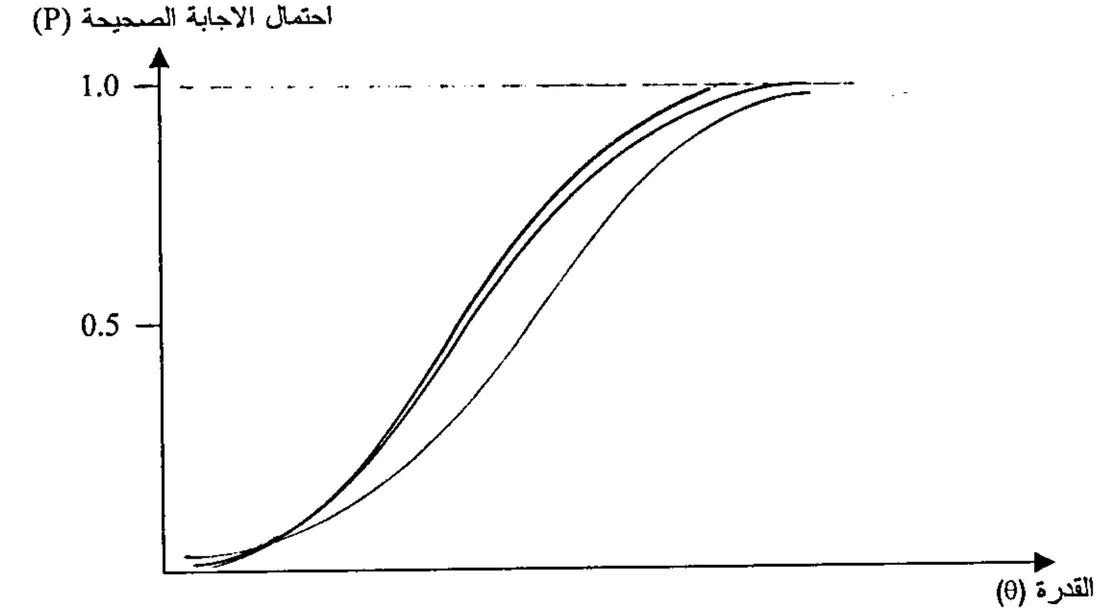
b_i = صعوبة الفقرة (i) .

e = الأساس اللوغاريتمي الطبيعي ويساوي (2.7183)

2- النموذج اللوغارتمي ثنائي المعالم Two-Parameter Logistic Model

ويسمى بنموذج لورد (Lord – Model) . يفترض هذا النموذج ان الفقرات تختلف في صعوبتها وتمييزها باختلاف المستويات المختلفة للقدرة ، ويفترض ايضا غياب التخمين (Ci= 0) .

وبســبب وجود معلم التمييز في هذا النموذج ، فان المنحنيات المميزة لفقرات الأختبار تتقاطع كما يوضحه شكل (2) :



الشكل رقم (2)

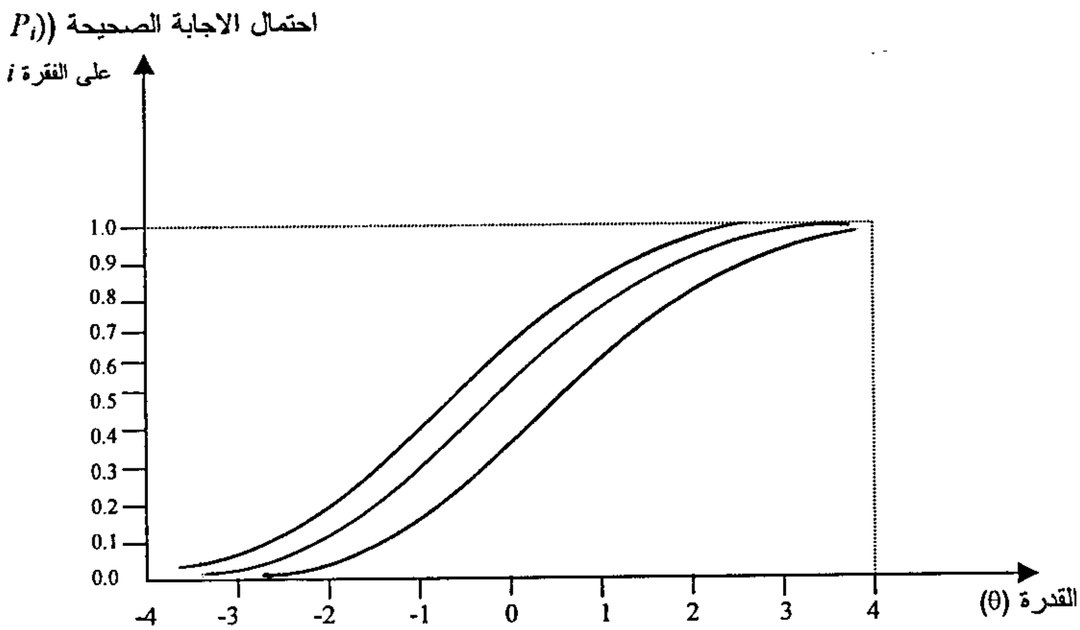
منحنى خصائص الفقرة ذو المعلمين لثلاث فقرات

ومعادلة هذا النموذج هي :

$$p(\theta) = \frac{e^{a(\theta - bi)}}{1 + e^{a(\theta - bi)}}$$

3- النموذج اللوغارتمي أحادي المعلم (نموذج راش) One-Parameter Logistic Model

ويسمى نموذج راش حيث يعتبرحالة خاصة من النموذجين السابقين ، ويتطلب تطبيقه افتراض تساوي مقادير التمييز لجميع فقرات الأختبار بين الأفراد ؛ أي أن المنحنيات متوازية ولاتتقاطع ، ولكنها تختلف عن بعضها في مستوى الصعوبة فقط مع غياب أثر التخمين كما يوضحه شكل (3) الآتي :



الشكل رقم (3)

منحنى خصائص الفقرة احادي المعلم لثلاث فقرات

ومعادلة نموذج راش هي :

$$P(\theta) = \frac{e^{(\theta - bi)}}{1 + e^{(\theta - bi)}}$$

علـــى اعتبار أن تمييز الفقرة الذي يفترضه النموذج ذا قيمة ثابتة تساوي (1) أي أن (a=1) .

تعتـــبر الـــنماذج الثلاثة السابقة الأكثر استعمالا في مجال البحث ضمن نظرية اســـــتجابة الفقــــرة ، ويـــــرى هامبلـــــتون وســــوامنثان (Hambleton and Swaminthan,1985) أن اختـــيار الـــنموذج يعـــتمد علـــى عـــدد الفقـــرات وعدد المفحوصين، وأن تكون الفقرات ملائمة للنموذج ، وهذا يتطلب مطابقة بيانات الفقرة مـــع افتراضات النموذج بحيث يكون ترتيب صعوبة الفقرة بين باقي الفقرات ثابتا ، وأن تنتمي الفقرة فعلا الى مجموعة فقرات الأختبار التي تقيس سمة واحدة فقط .

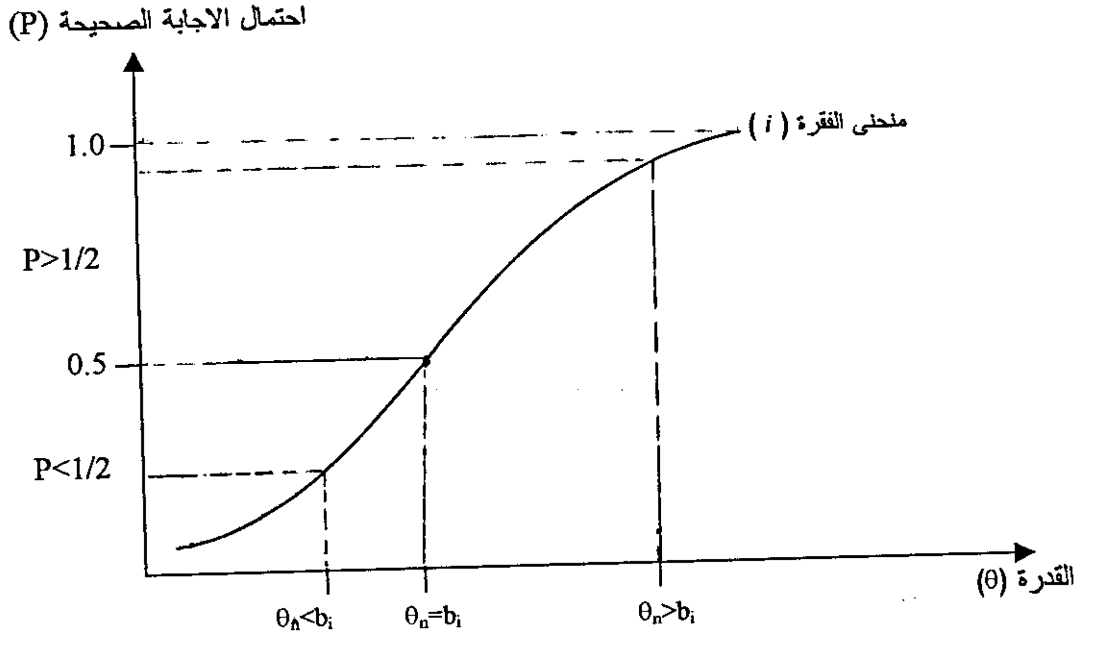
وفي هذه الدراسة تم استخدام نموذج راش ، لأن المواقف التعليمية تتطلب عددا قلـــيلا من الفقرات ومن المفحوصين ، إضافة إلى سهولة استخدامه في تحليل فقرات الاختـــبار إذا ما قورن بالنموذجين الآخرين ، كما يعد من أكثر النماذج التي أجريت علـــيها دراســـات للتحقق من خصائصه وجدواه في بناء الاختبارات ومواجهة أوجه القصـــور فـــي الـــنظرية الكلاســـيكية ، ويســـتخدم نموذج راش في تحليل الفقرات الاختـــبارية التـــي تكون الإجابة عنها إما صحيحة أو خاطئة ، بمعنى أن درجة كل فقرة تأخذ احدى قيمتين إما واحد أو صفر (علام ، 1995) .

لقد لخص كل من رايت وستون (Wright & Stone, 1979) العلاقة بين قدرة الفرد وصعوبة الفقرة عندما تتفاعلان لاستجرار الإجابة الصواب على فقرة ما على وذلك النحو الآتي :

عندما يستجيب الفرد (A) على الفقرة (i) تحدث الاستجابة XA_i وهي تساوي (واحد) عندما تكون الإجابة صوابا وتساوي (صفرا) عندما تكون الإجابة خاطئة.

ويـــتحكم فـــي إحـــداث هذه الإستجابة شرطان أساسيان هما : قدرة الفرد (θ) وصعوبة الفقرة (bi) .

ولمـــا كانت كل من (θ) و(bi) تقعان على متصل واحد "Continuum" فإن الفــرق $[\theta_A - bi]$ هو الصيغة الأكثر مناسبة للعلاقة بينهما والطريقة التي يؤثر بها الفرق $(\theta-bi)$ في الاستجابة (X_{Ai}) تكون احتمالية كما يوضـــحها الشكل (4) الآتي :



الشكل رقم (4)

العلاقة بين قدرة الفرد وصعوبة الفقرة

فـــإذا كان مستوى قدرة الفرد (θ_A) أكبر من صعوبة الفقرة (bi) كان الــفرق $(\theta-b_i)$ أكـــبر من الصفر وعندئذ يكون احتمال حدوث الاستجابة الصواب أكثر من (0.5) ، وإذا كانـــت (θ) أقـــل من (bi) أي أن مستوى قدرة الفرد أقل من صعوبة الفقـــرة كـــان الفرق $(\theta-b)$ أقل من الصفر، وعندئذ يكون احتمال حدوث الاستجابة الصواب أقل من (0.5) .

أما إذا كان الفرق $(\theta-b)$ مساويا للصفر عندئذ يكون احتمال حدوث الاستجابة الصـــواب مســـاويا لاحـــتمال حدوث الاستجابة الخطأ ويساوي كل منهما (0.5) ، وعليه يمكن تعريف صعوبة الفقرة بأنه مستوى القدرة اللازم كي يكون احتمال اجابة

الفـــرد اجابة صحيحة عن الفقرة عن الفقرة يساوي (0.5) ، وهذا يعني أن احتمال حدوث الاستجابة الصواب هي دالة " Function " لهذا الفرق بالرموز:

$$P_{Ai} = F(\theta\text{-}bi)$$

يلاحـــظ من خلال هذه الدالة أن احتمال الإجابة ينحصر بين القيمتين (صفر) و (واحـــد) ، في حين أن الفرق $(\theta\text{-}bi)$ يمكن أن يكون أي عدد حقيقي محصورا بيـــن -∞ و +∞ ، وقد أوضح العالم راش أن الدالة التي تحقق العلاقة السابقة هي الدالة اللوغارتيمية والتي صيغتها النهائية تكون على النحو الآتي :

$$Pi\ (\theta) = \frac{e^{(\theta - bi)}}{1 + e^{(\theta - bi)}}$$

ويسمى المقدار ($e^{\theta\text{-}b_i}$) مرجح النجاح "Odds of success " وهو نسبة احتمال النجاح $Pi(\theta)$ الى احتمال الفشل $1\text{-}Pi(\theta)$ حيث أن :

$$e^{\theta - bi} = \frac{Pi(\theta)}{1 - Pi(\theta)}$$

لقد أشار علام (1995) إلى أن نموذج راش يعتمـــد على قضيتين أساسيتين :

الأولـــى : أن احتمال الإجابة الصحيحة عن فقرة يزداد بزيادة قدرة الفرد .

والثانية : احتمال أن يجيب الفرد إجابة صحيحة عن فقرة سهلة أكبر من احتمال أن يجيب إجابة صحيحة عن فقرة صعبة .

ومـــن أهـــم ممـــيزات نمـــوذج راش هو ما يسمى بـــ" الموضوعية الخاصة "Specific Objectivity" (Fisher 1995) ، وتـــبدو هـــذه الموضـــــوعية مـــن ناحيتين :

الأولـــى : استقلال معلمة قدرة الفرد عن الفقرة المستخدمة حيث يحدد نموذج راش احـــتمال نجاح الفرد *(A)* في الإجابة عن الفقرة *(i)* بالمعادلة الآتية "معادلة راش" :

$$PAi = \frac{e\theta_A - bi}{1 + e\theta_A - bi}$$

حيث PAi ـ احتمال نجاح فرد قدرته (θ) على فقرة صعوبتها (b)

وأن e^{0-b} هو مرجح النجاح حيث :

$$e^{0\text{-}bi} = \frac{pAi}{1 - pAi}$$

وبأخذ لوغارتيم الطرفين ينتج ان :

$$\theta_A\text{-}b_i = ln\ \frac{PAi}{1 - PAi} \ldots \ldots \ldots (1)$$

وبالمثل في حالة استجابة فرد آخر *(u)* على الفقرة نفسها (i) فإن :

$$\theta_u\text{-}b_i = ln\ \frac{Pui}{1 - Pui} \ldots \ldots \ldots (2)$$

وبطـــرح المعادلتيـــن *(1,2)* مـــن بعضهما البعض فإن صعوبة الفقرة *(bi)* تحذف وينتج أن :

$$\theta_A\text{-}\theta_u = ln \frac{PAi}{1 - PAi} - ln\ \frac{Pui}{1 - Pui}$$

وهــذا يعنـــي أنــنا نستطيع أن نقارن بين قدرات الأفراد بطريقة مستقلة عن صعوبة الفقرات ، وعلى هذا فإن ما نصل إليه ليس القدرة المطلقة للفرد *(A)* وإنما بعـــده عـــن قدرة فرد آخر هو *(u)* وهذا الفرق يجعل الفرد (u) نقطة أصل تقاس مـــنها قـــدرة الفرد *(A)* ، وعليه فان تدريج الأفرادحسب قدراتهم على متصل السمة غير مرتبط بفقرات معينة (Item- Free) .

الثانــية : استقلال معلمة صعوبة الفقرة عن قدرة الفرد الذي يجيب عليها ، بالمثل إذا أجـاب الفـرد *(A)* علـى فقرتين من الفقرات *(k),(i)* فإننا نصل إلى المعادلتين الآتيتين :

$$\theta_A\text{-}b_i= ln\ \frac{PAi}{1-PAi}\ ....(3)$$

$$\theta_A\text{-}b_k = ln\ \frac{PAk}{1-PAk}\ ....(4)$$

وبطرح المعادلتين (3,4) من بعضهما البعض يحذف معلم القدرة *(θ)* وتبدو المقارنة بين الفقرتين *(k, i)* هي :

$$b_k\text{-}\ b_i= ln\ \frac{PAi}{1-PAi}\text{-}\ ln\ \frac{PAk}{1-PAk}$$

وعلى الرغم من اعتماد المقارنة بين صعوبتي فقرتين على إجابة فرد مناسب، إلا أن هـذه المقارنة لا تتأثر باستجابة أي من الأفراد ، وهذا ما نعنيه بأن المقارنة بيـن صـعوبة الفقرات تكون مستقلة عن الفرد (Person Free) حيث أن استجابة فرد لفقرتين ينبغي أن تؤدي المقارنة نفسها بين هاتين الفقرتين .

وعلـى هـذا فإن ما يتم التوصل إليه ليس الصعوبة المطلقة للفقرة *(i)* وانما بعدهـا عن صعوبة فقرة اخرى هي *(k)* وهذا الفرق يجعل من صعوبة الفقرة *(k)* نقطـة يقاس منها صعوبة الفقرة *(i)* ، وعليه فان تدريج الفقرات غير مرتبط بأفراد العينة "Sample- Free " .

ويذكـر رايـت (Wright ,1977) أن قدرة الفرد وصعوبة الفقرة تكونان على مقـياس مشـترك " Common Scale" لـه وحدات قياس محددة و موحدة تسمى وحـدات الترجـيح اللوغاريتمي" لوجيت Logits" ، ونقطة الصفر على تدريج هذا المقياس هي النقطة التي ترد إليها كل من معلم قدرة الفرد ومعلم صعوبة الفقرة .

واذا اعتـبرنا أن متوسط معاملات صعوبة الفقرات المقاسة هي نقطة صفر التدريج ، فإن هذه النقطة *(b=0)* تستخدم في تقديرات الأفراد .

وعند تعويض $(b=0)$ في المقدار($e^{\theta-b}$) الذي يعبر عن مرجح النجاح ، فان مرجح النجاح في هذه الحالة يساوي :

$$\text{مرجح النجاح} = e^{\theta}$$

وبأخذ لوغاريتم طرفي المعادلة السابقة عندئذ ، فإن اللوغاريتم الطبيعي لمرجح النجاح يساوي قدرة الفرد (θ) مقدرا باللوجيت .

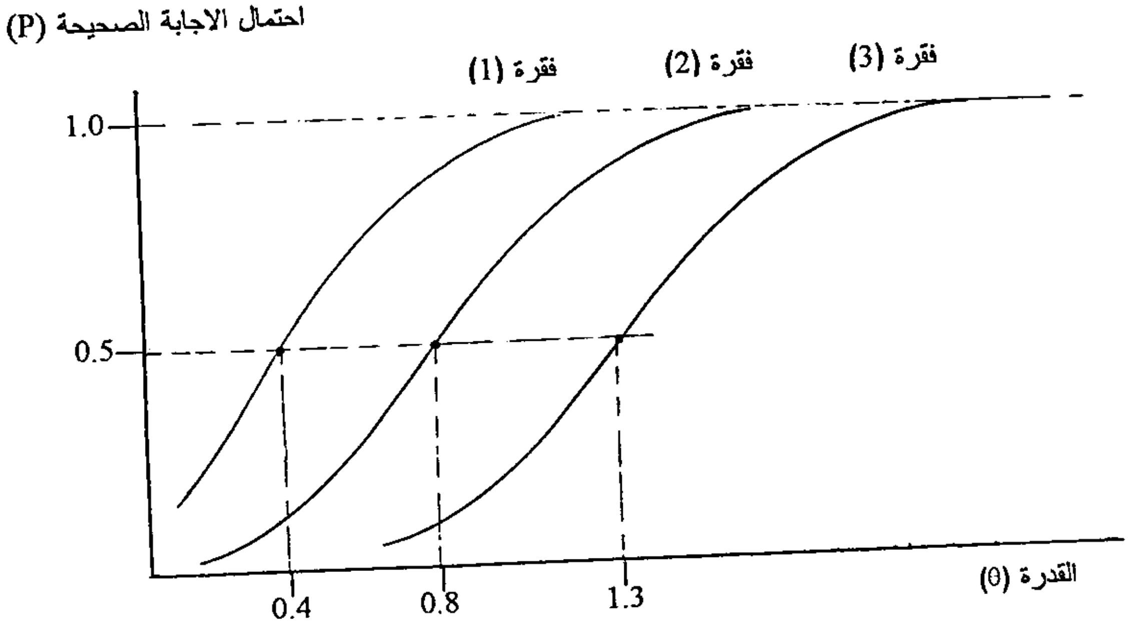
ومــن هنا عرّف كل من رايت وستون (Wright & Stone, 1999) قدرة الفرد بوحدة اللوجيت بأنها اللوغاريتيم الطبيعي لمرجح نجاح الفرد على الفقرات التي تعبّر نقطة صفر التدريج عن صعوبتها .

فعــندما تكــون قدرة فرد ما مساوية للوجيت واحد ، فهذا يعني أن اللوغاريتم الطبيعــي لمــرجح نجاح هذا الفرد على الفقرات التي تعبر نقطة صفر التدريج عن صــعوبتها يساوي واحد ، وبالتالي يمكن احتمال حدوث الاستجابة الصحيحة حسب علاقة راش :

$$P = \frac{e\theta - b}{1 + e\theta - b} = \frac{e\theta}{1 + e\theta} = \frac{2.72}{3.72} = 0.73$$

وهذا يعني إذا كانت قدرة فرد ما تساوي لوجيت واحد " One Logit" ، فإن احتمال حدوث الاستجابة الصحيحة له يساوي (0.73) .

اما صعوبة الفقرة حسب نظرية استجابة الفقرة ومن ضمنها نموذج راش فهي ليســت عدد الأفراد الذين يجيبون عنها إجابة صحيحة كما في النظرية الكلاسيكية ، وإنما هي النقطة على متصل السمة الكامنة التي نتوقع عندها أن يكون احتمال إجابة فــرد مــا صــحيحة عــن الفقــرة وبدون تخمين يساوي (0.5) . ويوضح بيكر (Baker,1985) ذلــك بحيــث لو فرضنا أن لدينا ثلاث فقرات يختلف موقع كل منها على متصل السمة ولها المستوى نفسه من التمييز فيلاحظ في هذه الحالة أنها تختلف فــي مقدار الصعوبة بحيث أن أسهلها هي الفقرة (1) وأصعبها هي الفقرة (3) كما هو مبين في الشكل التالي :



الشكل رقم (5)

منحنى خصائص الفقرة لثلاث فقرات مختلفة الصعوبة

فالفـــرد الذي تقع قدرته عند النقطة (0.4) على متصل السمة الكامنة يكون مقداراحـــتمال إجابـــته إجابة صحيحة عن الفقرة (1) يساوي (0.5) بينما لا يكون الحـــال كذلك عند محاولتة الإجابة عن أي من الفقرتين (2) و (3) وذلك لأن الفقرة (1) أقل صعوبة من كل من الفقرتين (2)، (3) بالنسبة لدرجته في السمة الكامنة .

ولـــتقدير كل من معلم الصعوبة الفقرة *(b)* ومعلم قدرة الفرد *(θ)* أشار رايت وآخـــرون(Wright , Mead & Bell., 1985) أنّ هـــناك عدة أساليب للتقدير ، منها مايعتمد على طريقة تسمىالارجحية العظمى" maximum Likelihood " ، ومنها ما يعتمد على طريقة تسمى احصاء بييز"Bayesian Statistic".

فـــي هذه الدراسة تم استخدام برنامج" Winsteps " وهو من أحدث البرامج المســتخدمة فـــي تحليل الفقرات حسب نموذج راش والذي طوعه للأستخدام رايت (Wright)مـــن جامعة شيكاغو بالمشاركة مع جون لينكر(Lincare) عام 1998 ، حيـــث يمتاز البرنامج باستخدامه طريقة الارجحية العظمى لتقدير كل من معلم قدرة الفرد ومعلم صعوبة الفقرة ضمن خطوتين :

الأولـــى تســمى الارجحــية العظمىغــير المشــروطة "Unconditional maximum Likelihood Procedure" واختصـــارها " UCON" ، و الثانـــية تسمى طريقة كوهين التقريبية ".Approximation" واختصارها " prox " ، بحيث قــبل الــبدء في استخدام هاتين الخطوتين يحذف كل فرد أخفق في حل جميع فقرات الاختبار (أي حصل على الدرجة صفر) ، لأن قدرة الفرد في هذه الحالة تعد أقل من مســتوى الاختــبار ، كمـــا يحذف كل فرد نجح في حل جميع فقرات الاختبار (أي حصــل علـــى الدرجـــة النهائية) لأن قدرته في هذه الحالة أعلى من مدى مستوى الاختــبار . وفـــي كــلا الحالين يكون هؤلاء الأفراد غير ملائمين أي غير مناسبين للإجابة عن هذا الاختبار .

كمـــا تحـــذف أيضا الفقرة التي يخفق جميع الأفراد في الإجابة عنها حيث تعد صـــعوبة الفقرة أعلى من مستوى قدرة الأفراد وكذلك الحال بالنسبة لكل فقرة يجيب عنها جميع الأفراد إجابة صحيحة حيث تعد صعوبتها أقل من مستوى قدرة الأفراد ، وتكون هذه الفقرات غير مناسبة لاستجابة الأفراد عليها .

بالأضـــافة لذلـــك فأن البرنامج يستخدم وحدة اللوجيت "Logit" لقياس كل من صـــعوبة الفقـــرة وقدرة الفرد والتي يمكن تغييرها خطيا باجراء التحويلات الخطية عليها . (Wright & Lincar, 1998) .

بعد الانتهاء من تقدير معالم كل من صعوبة الفقرة وقدرة الفرد ، فان الخطوة التالـــية فـــي نموذج راش إجراء تدريج " Calibration"على كل من صعوبة الفقرة *(b)* وقدرة الفرد *(θ)* ، وضحها لورد (Lord,1980) بأنها عملية وضع كلا المعلمين الصـــعوبة والقدرة على نفس التدريج بحيث يتم تحويل كل من قيم الصعوبة والقدرة على مقياس خطي وبنفس الوحدة وهي اللوجيت " Logit" باستخدام العلاقتين اللتين تم الحديث عنها في الموضوعية الخاصة لنموذج راش .

وبعــد تسليط الضوء على قدرة الفرد وصعوبة الفقرة في كل من النظريتين ، فأنه يمكن عقد مقارنة بينهما على النحو الآتي :

فمــن حيث العلاقات الرياضية فقد أشار ألن وين (Allen &Yen, 1979) الى أن الــنظرية الكلاســيكية تحدد العلاقة بين العلامة الملاحظة(X) والعلامة الحقيقية (T) ومقدار الخطأ العشوائي (E) بانها علاقة تجميعية خطية بسيطة :

$$T = X - E$$

أما في نظرية استجابة الفقرة وحسب نموذج راش فأ ن العلاقة تتحدد بين قدرة الفــرد ، واحــتمال اســتجابته عــن الفقرة بأنها علاقة غير خطية وانما لوغارتمية " Logarithmic " :

$$P = \frac{e^{\theta - b}}{1 + e^{\theta - b}}$$

وعليه يمكن الاستنتاج بأن العلاقة بين صعوبة الفقرة وقدرة الفرد (المفحوص) فــي نظرية استجابة الفقرة –حسب علاقة نموذج راش – هي علاقة واضحة ترتبط فــي اقتران واحد ومحددة بمنحنيات خصائص الفقرة ، بينما في النظرية الكلاسيكية العلاقة بينهما غير محددة ولا ترتبط باقتران واحد (Hambleton and Jones, 1993).

وفــي الــنظرية الكلاســيكية تعبّر علامة المفحوص عن القدرة ، وهي تقدير لعلامــته الحقيقية وتتدرج هذه التقديرات على متصل درجات الأختبار، اما درجات القــدرة حسب نظرية استجابة الفقرة فهي تتدرج نظريا على متصل القدرة من – ∞ الى +∞ وعمليا من –3 الى +3 (Crocker and Algena,1986).

أمــا من حيث الافتراضات فهناك اجماع بين علماء القياس على أن افتراضات النظرية الكلاسيكية ضعيفة "Weak Assumptions" ، بمعنى أن اي بيانات يمكن أن تحقق افتراضاتها وهذا سبب شيوعها ، أما افتراضات نظرية استجابة الفقرة فهي قويــة ولــيس من السهل لأي بيانات ان تحقق افتراضاتها مما حدّ بالتالي من شيوع تطبيقاتها (Hambleton and Swaminathan, 1985).

الدراسات السابقة

لــم يحظ موضوع مقارنة النظرية الكلاسيكية بنظرية استجابة الفقرة من خلال نموذج راش أو غيره من النماذج بدراسات كافية في البلاد العربية بشكل عام – في حــدود علــم الباحــث – ، وحــتى فــي الدراســات الأجنبية فان هناك نقصا في هذاالموضوع كما أشار فان في دراسته (Fan , 1998) .

والمقارنة المقصودة هنا هي المقارنة التطبيقية "Empirical" المستندة الى افتراضات كل من النظريتين باستخدام بيانات حقيقية أو بيانات تم توليدها عن طريق المحاكاة "Simulation" ، لذا ولغايات هذه الدراسة قام الباحث بالاطلاع على ما توفر من دراسات أجنبية وعربية مرتبطة بموضوع البحث مباشرة ، وقام بتصنيفها الى دراسات استخدمت بيانات حقيقية "Real Data " ، و أخرى استخدمت بيانات محاكاة "Simulation Data" وذلك على النحو الآتي:

أولا: الدراسات التي استخدمت بيانات حقيقية

1- الدراسات الاجنبية

فقــد أجــرى هوانــج (Hwang , 2002) دراســة هدفت الى المقارنة بين احصـــائيات الفقـــرة من خلال النظرية الكلاسيكية ونظرية استجابة الفقرة باستخدام بـــيانات تـــم الحصول عليها من استجابات أفراد عينة من المفحوصين عدد أفرادها (600) أجـــابوا عن (15) فقرة اختبار رياضيات للصف الثامن ، وقد تمت المقارنة بايجاد معاملات الارتباط لكل من صعوبة الفقرة وتمييزها وفق النظريتين باستخدام بـــرنامج "ITMAN " للنظـــرية الكلاسيكية وبرنـــامج "BILOG " لنظرية استجابة الفقرة .

أظهرت نتائج الدراسة أن افتراضات نظرية استجابة الفقرة قوية ، اذ ليس من السهل لأي بيانات أن تحققها وأن معاملات الارتباط لأحصائيات الفقرات أظهرت تشابها كبيرا فيما بينها وفق كل من النظريتين.

كما أجرى كل من كريسلب وتشن (Crislip & Chin, 2001) دراسة هدفت الى مناقشة استخدام النظرية الكلاسيكية ونموذج راش- احد نماذج نظرية استجابة الفقرة- في تحليل الفقرة وبناء الاختبار من حيث قوة وضعف هاتين النظريتين ، وتم تطبيقهما على اختبار مقاطعة هاواي للكفايات الأساسية "HSTEC" الذي يستخدم لتصنيف الطلبة بعد الصف العاشر ، وتكوّن الاختبار من (60) فقرة تم تطبيقه على عينة مكونة من (836) طالبا من طلبة الصف العاشر تم اختيارهم عشوائيا من سبع مدارس عبر مقاطعة هاواي .

وبعد التطبيق تم تحليل خصائص فقرات الاختبار من حيث معامل الصعوبة والتمييز على ضوء نتائج الطلبة استنادا لهاتين النظريتين . وأظهرت نتائج الدراسة بأن استخدام كلا النظريتين : الكلاسيكية ونظرية استجابة الفقرة خلال عملية بناء فقرات الأختبار ستزيد من ملائمة اداة القياس في تقديم بيانات يمكن الأعتماد عليها لعمل استنتاجات مقنعة عن مستويات الطلبة ، و عدم الاقتصار على خطوات نظرية واحدة فقط في تطوير أدوات القياس لسمة ما.

وأجــرت ســتيج (Stage,1999) دراســة هدفت الى مقارنة تحليل الفقرة بين نظــرية اســتجابة الفقرة والنظرية الكلاسيكية ، باستخدام اختبار القدرات للمدارس الســويدية "SweSat " الــذي يســتخدم في اختيار الطلبة لمرحلة التعليم العالي في السويد ويعقد مرتين في السنة ، علما بأن الاختبار ومنذ عام 1977 وهو يبنى حسب النظرية الكلاسيكية ، ثم في عام 1996 تم إعادة النظر في بناء الاختبار ليتوافق مع الــنظرية الحديــثة ، ويتكون الاختبار من (122) فقرة من نوع الاختيار من متعدد موزعة على مهارات متنوعة ، حيث تم تطبيق الاختبار على عينة حجمها (1500) مفحوصا .

أظهــرت نتــيجة الدراسة أن هناك توافقا بين الخصائص السيكومترية لفقرات الاختــبار بحيــث وصــل معامل الارتباط بين قيم معاملات صعوبة الفقرة حسب النظريتيــن (R= 0.92) ، لكن وجد ان الخصائص السيكومترية باستخدام النظرية الكلاســيكية يعــتمد على عينة المفحوصين بينما هذا لم يظهر عند استخدام نظرية استجابة الفقرة .

وقام اندرسون (Anderson, 1998) بدراسة هدفت إلى الكشف عن استخدام نظرية استجابة الفقرة في تحليل نتائج الاختبار باستخدام برنامج "Bilog" ومقارنتها بالنظرية الكلاسيكية ، واستندت الدراسة على استجابات واقعية من خلال اختبار الرياضيات الخاص بمقاطعة كولومبيا البريطانية والذي تم تطبيقة على عينة من طلبة الصف الثاني عشر عدد أفرادها (6147) طالبا أجابوا عن (50) فقرة من نوع الاختيار من متعدد ، حيث تم توليد اختبارين عن طريق أخذ الفقرات ذات الأرقام الفردية للأختبار الأول ، والفقرات ذات الأرقام الزوجية لفقرات الاختبار الثاني أي بواقع (25) فقرة لكل اختبار ، ثم تم تطبيق كل اختبار على حدة على العينة نفسها .

أظهرت النتائج أن معامل الارتباط بين علامات الطلبة على الاختبارين كانت كالآتي :

	الطريقة التقليدية CTT		طريقة IRT	
	test 1	test 2	test 1	test 2
	X_1	X_2	IRT_1	IRT_2
X_1	---	---	---	---
X_2	.783	---	---	---
IRT_1	.971	.775	---	---
IRT_2	.766	.977	.765	---

حيث ان (X1) ، (X2) تمثل علامات الطلبة – والتي تمثل قدراتهم– على الأختبارين باستخدام النظرية الكلاسيكية ، و(IRT1) ،(IRT2) تمثل قدرات الطلبة علــى الأختبارين باستخدام نظرية استجابة الفقرة ، من خلال الجدول اظهرت نتائج الدراســة ان معاملات الأرتباط بين قدرات الطلبة باستخدام نفس مجموعة الفقرات مــن خلال النظريتين(X1:IRT1,and X2:IRT2) تقريبا تساوي(0.97) ، بينما معــاملات الأرتــباط بين قدرات الطلبة ماخوذة من مجموعات مختلفة من الفقرات ولكــن باستخدام اسلوب التحليل نفسه ((X1:X2 and IRT1:IRT2 تقريبا تساوي (0.77) ، وهذا يشير الى أن النتائج متقاربة كثيرا بين النظريتين .

أمـــا دراســـة فـــان (Fan,1998) فقد هدفت إلى مقارنة إلى أي مدى تختلف النظرية الكلاسيكية عن نظرية استجابة الفقرة في تقدير قدرات المفحوصين ، وتقدير معاملات صعوبة الفقرة ، و تقدير معاملات تمييز الفقرة .

حيث استخدم الباحث بيانات حقيقية من اختبار تكساس للمهارات الأكاديمية " TAAS " الـــذي طبق في عام 1992 على طلبة الصف الحادي عشر من خلال ثـــلاث عينات يبلغ عدد أفراد كل عينة (1000) طالب ، أجابوا عن (40) فقرة من نـــوع الاختـــيار مـــن متعدد ،جرى تحليل نتائج الاختبار حسب النظرية الكلاسيكية ونظرية استجابة الفقرة بنماذجها الثلاث باستخدام برنامج "Bilog ".

أظهـــرت النـــتائج أن تقدير قدرات الأفراد من خلال النظرية الكلاسيكية كانت مشابهة " Comparable" إلى حد كبير مع التقديرات من خلال نظرية استجابة الفقرة بـــنماذجها الثلاثة باستخدام معاملات الارتباط ، كما وجد أن معاملات صعوبة الفقرة مـــن خلال النظرية الكلاسيكية تتشابه مع التقديرات من خلال نظرية استجابة الفقرة باســـتخدام معـــاملات الارتـــباط ، وأن معاملات تمييز الفقرات من خلال النظرية الكلاسيكية تتشابه بصورة أقل من تقديرات نظرية استجابة الفقرة باستخدام معاملات الارتـــباط . ولكـــن ماميّز نظرية استجابة الفقرة عن النظرية الكلاسيكية فهو ثبات احصائيات الفقرة خلال العينات التي تم استخدامها.

واســـتنتج أن دراسته لم تدعم القول بوجود أفضلية لنظرية استجابة الفقرة من خـــلال اطارهـــا النظري ، بل لابد من اجراء المزيد من الدراسات التجريبية حول المقارنة بين النظريتين .

وأمـــا دراسة دوجلاس وفريزر (Douglass and Frazier,1979) (المشار اليها فـــي جمحـــاوي ، 2000) فقد هدفت الى مقارنة اجراءات تحليل الفقرات بالنظرية الكلاســـيكية ونظرية السمات الكامنة ، حيث جمعت استجابات (373) فردا غالبيتهم مـــن الطلبة الجامعيين البيض الذين أكملوا مقياس الكشف عن الكحوليين " Mac Andrew"المكـــون مـــن 49 فقرة و اختبار"Khavari Alcohol" المكون من 12 فقـــرة ، وهو مقياس لتحديد كمية استهلاك الكحول . حللت الاستجابات على مقياس الكشـــف عـــن الكحوليـــين "Mac Andrew" باستخدام كل من النظرية الكلاسيكية

ونمـــوذج راش – احد نماذج نظرية السمات الكامنة – ، وقورن المقياسان الناتجان مـــن كـــلا الاسلوبين على مستوى الفقرات المنتقاة وارتباطها بالاختبار الاصل وهو الاستهلاك الحقيقي للكحول.

أظهـــرت نـــتائج الدراســـة وجـــود فـــروق حقيقية بين الطريقتين – النظرية الكلاسيكية ونموذج راش – على مستوى اختيار الفقرات ، فبعض الفقرات التي تم انتقاؤها باستخدام نموذج راش تم استبعادها باستخدام النظرية الكلاسيكية وبالعكس ، وكذلك فان الفقرات التي تم انتقاؤها بكلا الطريقتين سلكت سلوكا مختلفا في علاقتها مـــع المقياس الاصل الذي اعتبر كمحك وهو (الاستهلاك الحقيقي للكحول) . وبناء على هذه النتائج فقد أوصت الدراسة الى اعتبار التحليل باستخدام النظرية الكلاسيكية الخيار الامثل نظرا لسهولة اجراء التحليل بها.

وأجرى هاشوي (Hashway,1977) دراسة هدفت الى الإجابة عن الأسئلة التالية:

هل يختلف تدريج المفحوصين بالأختبارات المطورة وفق نموذج راش عنه وفق النظرية التقليدية ؟ ، و هل توجد اختلافات سيكومترية جوهرية بين الفقرات المنتقاة باستخدام نموذج راش والمنتقاة باستخدام الطرق التقليدية ؟

وللإجابة عن أسئلة الدراسة تم اختيار عينة عشوائية من (1325) طالبا من الصف السادس في ايرلندا ، وتم تطبيق اختبار محكي المرجع في الرياضيات على هذه العينة ، وكان الاختبار مؤلفا من (155) فقرة تضمنت (55) هدفا ، وجمعت البيانات على مستوى الفقرة والهدف.

حسبت خصائص الفقرات وفق النظرية التقليدية ونموذج راش وبينت الدراسة الى أن كلا من نموذج راش والنظرية قد سلكا سلوكا مختلفا عن الآخر في تدريج المفحوصين وترتيبهم . كما قورنت صعوبة الفقرات المنتقاة باستخدام نموذج راش بالفقرات المنتقاة باستخدام النظرية التقليدية ووجد أن عدد المفحوصين الذين أجابوا اجابة صحيحة حسب نموذج راش لم يختلف اختلافا كثيرا عن عدد أولئك الذين اجابوا اجابة صحيحة حسب النظرية التقليدية مما يعني أن الفقرات المنتقاة وفق نموذج راش والنظرية التقليدية كانت متقاربة .

2- :الدراسات العربية

لقـد تـم الاطـلاع على دراستين باللغة العربية ذاتي صلة بموضوع البحث الحالـي مباشـرة أستخدمتا بيانات حقيقية ، ومن الجدير ذكره أن هاتين الدراستين محليتان لباحثين اردنيين وهما: دراسة جمحاوي (2000) ، ودراسةعودة (1992) .

أما دراسة جمحاوي فقد كان عنوانها (مقارنة خصائص الفقرات وفق النظرية التقليدية ونظرية استجابة الفقرة في مقياس للقدرة الرياضية) ، حيث هدفت هذه الدراسة الى المقارنة بين النظرية التقليدية والنظرية الحديثة في القياس، في مقياس للقدرة الرياضية والمطور وفق المؤشرات التقليدية ليوائم طلاب البيئة الاردنية وبالتحديد الصفوف الثامن والتاسع والعاشر . حيث قورنت خصائص فقرات هذا المقياس وعدده (39) فقرة وفق النظريتين ، وتركزت المقارنة في مدى اتفاق النظريتين في انتقاء الفقرات من حيث المحتوى والعدد ، وفي مدى اتفاقهما في تقدير صعوبة الفقرات وتمييزها ايضا ، وكذلك في مدى اتفاقهما في توزيع القدرة للمفحوصين ، وفي ثبات المقاييس الناتجة من استخدام كل نظرية .

تكـون مجـتمع الدراسـة من طلبة الصف التاسع في منطقة اربد ، أما عينة الدراسـة فقد تكونت من (1061) طالبا وطالبة من طلبة الصف التاسع في مديريتي اربـد الأولـى والثانية . وقد طلب من أفراد العينة الاجابة عن فقرات المقياس في ظـروف صـفية عاديـة ، وحللـت اجابات افراد الدراسة باستخدام برامج التحليل "ASCAL" و "RASCAL " للتحلـيل بـنماذج الـنظرية الحديـثة وبرنامجـي " ITMAN " و "SPSS" للتحلـيل بالـنظرية التقلـيدية ، وذلـك للحصول على الأحصائيات الخاصة بالفقرات وبالأفراد .

أظهرت نتائج الدراسة الى وجود اتفاق عالي بين النظريتين في تقدير صعوبة وتمييز الفقرات، وكذلك فقد أظهرت النظريتان اتفاقاً عالياً في توزيع القدرة للمفحوصين ، وبينت النتائج أيضاً أن عدد الفقرات التي تم اختيارها وفق المؤشرات التقليدية كان (35) فقرة مقابل (20) فقرة انسجمت ونموذج راش و(35) فقرة انسجمت ونموذج المعلمتين، و(38) فقرة انسجمت ونموذج المعالم الثلاثة من اصل

(39) فقرة، الأمر الذي يؤكد أن عدد الفقرات المنسجمة مع النموذج يزداد بزيادة عدد معالمه.

كما ويجب ملاحظة أن قيم معاملات التمييز والصعوبة التقليدية العالية لاتحقق أو لا تضمن تطابق الفقرات المختارة مع أحد نماذج السمات الكامنة، كما أن انخفاض معاملات التميز والصعوبة التقليدية لبعض الفقرات لايعني عدم اتفاقها مع أحد نماذج السمات الكامنة، كما وأن الزيادة في عدد الفقرات المنسجمة مع النموذج لا يوفر نسبة اتفاق عالية مع المؤ شرات التقليدية في انتقاء الفقرات.

وأظهرت النتائج أيضا أن سلوك نموذج المعلمتين كان الأكثر اتفاقاً دائماً والمؤشرات التقليدية من النماذج الأخرى؛ والسبب في ذلك- كما اشارت الباحثة – ربما لأنه الأكثر شبهاً بالمؤشرات التقليدية في اختيار الفقرات، الأمر الذي ربما جعله الأكثر اتفاقاً معها في تقدير قدرات الفحوصين وجعل الفرق بين معامل ثبات المقياس الناتج منه ومعامل ثبات المقياس الناتج من المؤشرات التقليدية غير دال إحصائياً.

وأما دراسة عودة (المشار اليها في جمحاوي ،2000) فقد هدفت الدراسة الى الكشف عن مدى التوافق بين نموذج راش ذي المعلم الواحد والمؤشرات الأحصائية التقليدية في اختيار فقرات اتجاهات المعلمين نحو الامتحانات المدرسية من حيث عدد ونوع الفقرات في صورتي المقياس ، ومدى تغطية الفقرات لمجال الاتجاه ومعاملات الثبات .

حللت نتائج اجابات (458) معلما على (42) فقرة بثلاثة مؤشرات احصائية تقليدية ، كما حللت باستخدام برنامج خاص بنموذج راش قبل وبعد فرز أفراد العينة الذين انسجمت اجاباتهم مع النموذج.

أظهرت نتائج الدراسة وجود درجة عالية من التوافق في عدد الفقرات ومعامل الثبات للمقياس على عينة الافراد الذين انسجمت اجاباتهم مع النموذج ، الا أن اختيار الفقرة على أساس معامل الارتباط لايعني بالضرورة انسجامها مع النموذج في مقياس يفترض أنه أحادي البعد .

ثانيا :الدراسات التي استخدمت بيانات محاكاة

لاتوجد دراسات في اللغة العربية – حسب اطلاع الباحث –استخدمت بيانات محاكاة في المقارنة بين النظرية الكلاسيكية ونظرية استجابة الفقرة أو أحد نماذجها ، وأنما وجد الباحث دراستين اجنبيتين هما : دراسة ماكدونالد وبونونن (Macdonald & Paunonen,2002) ، ودراسة تروسكوسكي (Truskosky,1999).

فقد اجرى ماكدونالد وبونونن (Macdonald & paunonen,2002) دراسة هدفت الأجابة عن الأسئلة الآتية:

كيف تتشابه احصائيات الفرد/ الفقرة باستخدام بيانات تم توليدها بالمحاكاة في ضوء النظريتين : الكلاسيكية واستجابة الفقرة ؟ .

كيف تتباين احصائيات الفقرة باستخدام كلا النظريتين عبر عينات المفحوصين ؟.

ما دقة تقدير احصائيات كل من الفرد والفقرة من خلال النظرية الكلاسيكية ونظرية استجابة الفقرة ؟ .

استخدم الباحثان طريقة مونت كارلو"Monte Carlo" في توليد بيانات مئة اختبار عن طريق المحاكاة "Simulation" تختلف في عدد فقراتها (20،40،60) فقرة ، وطبقت هذه الأختبارات على عينتين تم توليدهما عشوائيا وبلغ عدد أفراد كل عينة (1000) فرد .

وللأجابة عن أسئلة الدراسة حسبت أحصائيات كل فقرة : معامل صعوبتها وتمييزها ، وقدرة الفرد المفحوص وفق النظريتين : الكلاسيكية باستخدام برنامج "SPSS" ، ونظرية استجابة الفقرة باستخدام برنامج "PARASCALE" عبر النموذجين : ذي المعلمة الواحدة (نموذج راش) والنموذج ذي المعلمتين ، ثم تم ايجاد معاملات الارتباط بين قيم قدرات الافراد حسب النظرية الكلاسيكية ونموذج راش ، ومعاملات الارتباط بين قيم معاملات الصعوبة حسب النظرية الكلاسيكية ونموذج راش ومقارنة هذه القيم عبر عينات الدراسة .

أظهرت نتائج الدراسة أن معامل الارتباط بين قيم قدرات الافراد من خلال النظريتين لاتقل عن (97.0) ، وبيّنت الدراسة أن الارتباط العالي هذا يشير الى ان القرارات حول مستويات المفحوصين تتشابه كثيرا بغض النظرعن الاطار النظري الذي أخذ منه القياس سواء النظرية الكلاسيكية ام نظرية استجابة الفقرة .

أما بالنسبة لاحصائيات الفقرة : صعوبة الفقرات وتمييزها فبيّنت الدراسة أنها كانت ذات استقرار عالي بين العينات العشوائية للمفحوصين في نظرية استجابة الفقرة وبدرجة أقل في النظرية الكلاسيكية . أما دقة التقديرات لكل من صعوبة الفقرة وقدرة الفرد من خلال النظريتين فكانت ايضا عالية ، أما بالنسبة لتقديرات تمييز الفقرة فبينت الدراسة انها كانت عالية من خلال نظرية استجابة الفقرة عبر جميع عينات الدراسة .

أما دراسة تروسكوسكي (Truskosky,1999) فقد هدفت الى اختبار مدى دقة وثبات تقديرات كل من النظرية الكلاسيكية ونظرية استجابة الفقرة بمعالمها الثلاثة من خلال حجوم عينات مختلفة(1000، 750، 500 ، 300، 150، 75) تم اختبارهم على (60) فقرة باستخدام طريقة المحاكاة في توليد البيانات .

تم الحصول على تقديرات معالم نظرية استجابة الفقرة – صعوبة الفقرات وقدرات الافراد – باستخدام طريقة الارجحية الهامشية القصوى "Marginal Maximum Likelihood" ، وتقديرات معالم النظرية الكلاسيكية – صعوبة الفقرات وقدرات الافراد – عن طريق نسبة الاجابة الصحيحة والعلامة الكلية على التوالي . اظهرت نتائج الدراسة ان التقديرات لمعالم كل من النظريتين تكون متقاربة عندما يكون حجم العينة لايقل عن (500) فردا ، كما اظهرت نتائج الدراسة ان تقديرات نموذج نظرية استجابة الفقرة ذي المعلمتين وتقديرات معالم النظرية الكلاسيكية يمكن الحصول عليها من حجم عينات متكافئة مما يدحض الفرض القائل بأن حجم عينات نظرية استجابة الفقرة يجب ان تكون اكبر بكثير من حجوم عينات النظرية الكلاسيكية عند المقارنة .

من خلال ماسبق من دراسات يمكن الاستنتاج ان الدراسات السابقة قد ركزت على المقارنة بين النظريتين الكلاسيكية ونظرية استجابة الفقرة من خلال نوعين من البيانات :

الأول : دراسات اعتمدت على بيانات حقيقية لأختبارات عامة تهم شريحة واسعة من الطلبة مثل اختبار مقاطعة هاواي للكفايات الأساسية "HSTEC" كما في دراسة كريسلب وتشن (Crislip &Chin,2001) واختبار القدرات للمدارس السويدية "Swe Sat" كما في دراسة ستيج (Stage,1999) ، واختبار تكساس للمهارات الأكاديمية "TAAS" كما في دراسة فان (Fan,1998) ، ودراسة جمحاوي (2000) ،لأن هذه الأختبارات تطبق على عينات كبيرة حيث تراوح عدد أفراد عينات الدراسات السابقة مابين 836 الى 1500 مفحوصا ، وهذا يساعد في دقة المقارنة وخاصة اذا ما علمنا ان نظرية استجابة الفقرة تتطلب عينات كبيرة .

الثاني : دراسات اعتمدت على بيانات تم توليدها بطريقة المحاكاة "Simulation" مثل دراسة ماكدونالد وبونونن (Macdonald & Paunonen, 2002) ، ودراسة تروسكوسكي (Truskosky,1999) وذلك باستخدام برامج كمبيوتر معينة من اجل ضمان الحصول على اختبارات عديدة وتطبيقها على عينات عشوائية كبيرة ، لأنه – وكما اشار فان (Fan,1998) في دراسته – أن المقارنة بين النظرية الكلاسيكية ونظرية استجابة الفقرة من خلال بيانات حقيقية قد تكون مقيدة بطبيعة هذه الفقرات وبظروف تطبيق الاختبار ، لذا من المستحسن في هذه الحالة استخدام عدد كبير من الأختبارات والعينات والتي يتم توليدها عن طريق المحاكاة للوصول الى نتيجة محددة حول المقارنة بين النظرية الكلاسيكية ونظرية استجابة الفقرة.

ان وجود عيوب في النظرية الكلاسيكية جعلت بعض المؤسسات المعنية ببناء الاختبارات تتجه في بناء اختباراتها نحو نظرية استجابة الفقرة كما أشارت ستيج (Stage,1999) في دراستها ، حيث أن اختبار القدرات للمدارس السويدية (SweSat) ومنذ عام 1996 أصبح يبنى وفق نظرية استجابة الفقرة.

انطلقت الدراسات السابقة من الأطار النظري لكل من النظريتين : الكلاسيكية ونظرية استجابة الفقرة حيث اشارت الدراسات - كما في دراسة فان (Fan,1998)- الـــى ان الاطار النظري واضح لكل من النظريتين . أما المعرفة التطبيقية فلاتزال ناقصـــة لـــذا لابـــد مـــن التوجه الى دراسات تجريبية "Empirical" للتحـــقق من المعرفة النظرية لمزايا كل من النظريتين.

كمـــا أشـــارت جمـــيع الدراسات الى التقارب الكبير في تقديرات كل من قيم صـــعوبة الفقـــرات وقدرات الأفراد من خلال المقارنة بين النظريتين الا ان نظرية اســـتجابة الفقـــرة تمـــتاز بأنهـــا تحـــرر قيم صعوبة الفقرات من عينة المفحوصين "Sample- Free" ، كمـــا تحـــرر قـــيم قدرات الأفراد من عينة الفقرات التي تم اختبارهم بها "Item- Free" مما يحقق الموضوعية في القياس.

الفصل الثالث

المنهجية والإجراءات

يتضمن هذا الفصل وصفا لمجتمع الدراسة وعينتها ، وطريقة اختيارها ، وطريقة تطبيق فقرات الاختبار عليها ، بالإضافة إلى وصف لإجراءات الدراسة والطرق الإحصائية التي استخدمت في معالجة البيانات وتحليلها لإجراء المقارنة لكل من قدرة الفرد وصعوبة الفقرة من خلال نموذج راش والنظرية الكلاسيكية .

مجتمع الدراسة

يتكون مجتمع الدراسة من جميع طلبة الصف التاسع الأساسي في مديريات التربية والتعليم والثقافة العسكرية والتعليم الخاص ووكالة الغوث في الأردن والبالغ عددهم تقريبا (28065) طالبا وطالبة خلال العام الدراسي 2001/2000م .

عينة الدراسة

تـــم اختيار عينة عشوائية من نوع العنقودية المتعددة نسبتها (5%) من مجتمع الدراســـة حيـــث بلـــغ عددهم (5613) طالبا وطالبة موزعين على (164) مدرسة مناصفة بحسب الجنس ، وشملت مجموعة من مدارس المدن والريف منها : (112) فـــي المـــدن ، (52) مدرسة في الريف ، وبينت هلن وآخرون (Hulin,Lissak,& Drasgow 1982) المشـــار إليهم في (Crocker and Algena, 1986) أن حجم العينة المناســـب لـــتقدير إحصـــائيات الفقرة من خلال نظرية استجابة الفقرة يعتمد على الـــنموذج المســـتخدم بحيث يزداد حجم العينة بازدياد عدد معالم النموذج ، ولكن في الغالب يجب أن لا يقل حجم العينة عن (1000) للنموذج ذي المعالم الثلاث .

والجدول رقم (2) يبين توزيع أفراد عينة الدراسة حسب المديرية وعدد الطلبة وعدد المدارس والنمط السكاني لها .

جدول رقم (2)

توزيع أفراد عينة الدراسة حسب المديرية والجنس وعدد المدارس

المديرية	عدد مدارس الذكور	عدد الطلاب	عدد مدارس الاناث	عدد الطالبات	المجموع		النمط السكاني	
					عدد المدارس	عدد الطلبة	مدارس مدن	مدارس قرى
عمان/1	7	280	9	360	16	640	16	0
عمان/2	5	191	5	178	10	369	10	0
عمان/3	2	58	2	50	4	108	4	0
عمان/4	4	160	3	120	7	280	7	0
مادبا	2	61	2	56	4	117	2	2
الزرقاء	6	140	7	261	13	401	9	4
الرصيفة	2	80	2	80	4	160	4	0
السلط	2	78	2	80	4	158	2	2
عين الباشا	1	32	1	26	2	58	0	2
دير علا	1	22	1	20	2	42	0	2
الشونة الشمالية	1	9	1	19	2	28	1	1
إربد الأولى	6	214	5	179	11	393	6	5
إربد الثانية	2	80	2	78	4	158	1	3
الكورة	2	60	2	55	4	115	1	3
بني كنانة	1	40	1	40	2	80	0	2
الأغوار الشمالية	1	40	1	30	2	70	0	2
الرمثا	2	55	3	93	5	148	3	2
جرش	2	68	2	80	4	148	2	2
عجلون	3	104	2	67	5	171	2	3
المفرق	2	80	2	80	4	160	2	2
البادية الشمالية	2	44	2	35	4	79	0	4
الكرك	2	80	2	57	4	137	3	1
الأغوار الجنوبية	2	33	2	38	4	71	1	3
المزار الجنوبي	1	40	2	50	3	90	2	1
القصر	1	20	1	19	2	39	0	2
الطفيلة	2	61	3	105	5	166	3	2
معان	1	40	2	55	3	95	2	1
العقبة	2	80	1	40	3	120	2	1
الثقافة العسكرية	1	40	1	24	2	64	2	0
وكالة الغوث	9	360	8	320	17	680	17	0
التعليم الخاص	5	163	3	105	8	268	8	0
المجموع	**82**	**2813**	**82**	**2800**	**164**	**5613**	**112**	**52**

أداة الدراسة

هدفت هذه الدراسة الى المقارنة بين صعوبة الفقرة وقدرة الفرد في ضوء كل من النظرية الكلاسيكية و نموذج راش في القياس ، ولتحقيق هذا الهدف فقد تم استخدام نتائج الطلبة على الاختبار الوطني في مبحث العلوم للصف التاسع الأساسي الذي تم تطبيقه على عينة الدراسة من قبل وزارة التربية والتعليم للعام الدراسي 2001/2000م . وتكون الاختبار الوطني ملحق رقم (1) من نوعين من الأسئلة :

النوع الأول : أسئلة الاختيار من متعدد وعددها (20) فقرة لكل منها أربعة بدائل واحد منها فقط صحيح ، وهي التي تم استخدامها لتحقيق أهداف الدراسة .

النوع الثاني : أسئلة مقالية وعددها ثمانية .

روعي في تطبيق إجراءات الاختبار المنهجية العلمية في بناء الاختبارات من حيث : تحليل محتوى الكتب المدرسية لمبحث العلوم من الصف الأول الأساسي إلى الصف التاسع الأساسي ، وإعداد جدول مواصفات ، وبناء الفقرات ، وتجريبها ، وتحليلها إحصائيا ، واختيار فقرات الاختبار ، وقام بهذا العمل فريق من المختصين كان الباحث أحد أعضائه ويبين الملحق رقم (2) أهداف الاختبار وخطوات بنائه وإجراءات تطبيقه .

إجراءات الدراسة

يتلخص الأسلوب الذي تم اتباعه لتحقيق أهداف الدراسة بالخطوات الآتية :

أولا : جمع البيانات :

تم الحصول على نتائج الطلبة على فقرات الاختبار الوطني في مبحث العلوم للصف التاسع الأساسي من مديرية الامتحانات في وزارة التربية والتعليم ، وتبين أن عدد الطلبة الذين خضعوا لهذا الاختبار بلغ (5520) طالبا وطالبة من أصل عينة الدراسة التي عدد أفرادها (5613) طالبا وطالبة .

تم عزل استجابات الطلبة على اختبار فقرات الاختيار من متعدد للاختبار الوطني البالغة (20) فقرة والمصححة ثنائيا (0 ، 1) وتخزينها على قرص كمبيوتر مرن "Floppy disk" وحفظها في ملف "wordpad" وذلك للتعامل معها باستخدام الرزمة الإحصائية "winsteps" الخاصة بنموذج راش والرزمة الإحصائية "SPSS" حسب النظرية الكلاسيكية .

ثانيا : التحليل العاملي :

قبل المضي في معالجة البيانات التي تنتج عن استخدام الرزمة الإحصائية (winsteps) تم إدخال البيانات إلى ذاكرة الحاسوب لفحص افتراض أحادية البعد للسمة المقيسة (قدرة طلبة الصف التاسع على الإجابة عن فقرات اختبار تحصيلي في مبحث العلوم) ، باعتبار أن هذا الافتراض هو أحد الافتراضات الأساسية التي يجب توافرها في البيانات عند استخدام نموذج راش ، اذ لا جدوى من الدراسة إن لم يتحقق هذا الافتراض ، وتم ذلك بالطريقتين الآتيتين :

الطــريقة الأولــى : عمل تحليل عاملي للبيانات الناتجة بتحليل المكونات الأساسية " Principal Component Analysis" وملاحظة قيم الجذر الكامن " Eigen Value " ونسب التباين المفسر لكل من العاملين الأول والثاني ، حيث في هذا النوع من التحليل تعتمد أحادية البعد إذا كانت نسبة الجذر الكامن للعامل الأول إلــى الجــذر الكامــن للعــامل الثانـــي تكون نسبة كبيرة لا تقل عن (2) . (Hambeleton & Swaminathan, 1985) .

الطريقة الثانية : باستخدام محك كاتل " Cattell " ويطلق عليه اختبار فرز العوامل " Scree Test" وتقوم فكرته على رسم محورين متعامدين : الأفقي وعليه عدد العوامل ، والعمودي الذي يعبر عن الجذر الكامن حيث تتناقص قيم الجذر الكامن بشكل كبير في العوامل الأولى التي تشير إلى وجود عامل عام مما يؤكد أحادية البعد .

التعريفات الأجرائية

قدرة الفرد في النظرية الكلاسيكية

هي العلامة المقدرة التي يحصل عليها الفرد عند إجابته تحريريا أو شفويا على فقرات اختبار يقيس سمة معينة .

قدرة الفرد في نموذج راش

هي اللوغاريتم الطبيعي لمرجح نجاح الفرد على الفقرات التي تعبّر نقطة صفر التدريج عن صعوبتها (Wright & Stone, 1979) .

معامل صعوبة الفقرة في النظرية الكلاسيكية

هـــو نســـبة عـــدد الذيـــن أجـــابوا اجابة صحيحة على الفقرة الى عدد المفحوصين عن تلك الفقرة .

معامل صعوبة الفقرة في نموذج راش

هـــو نقطـــة على متصل السمة التي نتوقع عندها أن يكون احتمال إجابة الفرد المفحوص عن الفقرة إجابة صحيحة بدون تخمين مساويا لـــ 0.5 .

الاختبار الوطني

اختـــبـــار تحصيلي تم اعداده من قبل وزارة التربية والتعليم في الاردن للعام الدراسي 2001/2000 بهدف قياس مدى امتلاك طلبة الصف التاسع للمهارات الأساسية في مبحث العلوم .

المعالجات الإحصائية

للإجابة عن أسئلة الدراسة أجريت المعالجات الإحصائية الآتية :

1 - إحصاء وصفي " Descriptive " لأداء عينة الدراسة للتحقق من عدد أفراد العينة التي سوف تدخل في عملية التحليل الإحصائي حسب كل من نموذج راش والنظرية الكلاسيكية وكذلك للتحقق من عدد الفقرات والوسط الحسابي للعلامات الكلية على الاختبار وأعلى علامة وأدنى علامة والانحراف المعياري للعلامات الكلية وشكل توزيع العلامات على الاختبار ومدى مماثلتها للتوزيع الطبيعي "Normal Distribution " .

2 - إيجاد قيم قدرات الطلبة المفحوصين وقيم صعوبة الفقرات حسب النظرية الكلاسيكية بتحليل البيانات باستخدام الرزمة الإحصائية " SPSS "حيث تم إيجاد العلامة الكلية لكل مفحوص " Total Score " والتي تعد تقديرا لقدرته كما أشار بذلك هامبلتون وآخرون (Hambeleton, Swaminthan, 1991 .)& Rogers . أما معامل صعوبة كل فقرة فتم إيجادها من نسبة المفحوصين الذين أجابوا عن الفقرة إجابة صحيحة .

وأما قيم قدرات الطلبة المفحوصين وصعوبة الفقرات حسب نموذج راش فقد تم استخدام الرزمة الإحصائية (Winsteps) في إيجاد هذه القيم ، حيث لم يتم حذف أي طالب من الطلبة المفحوصين على اعتبار أنه لا يوجد أي طالب فشل في الإجابة عن جميع الفقرات (أي حصل على الدرجة صفر) ، و لا يوجد أي طالب أجاب عن جميع الفقرات (أي حصل على الدرجة الكلية الكاملة) ، حيث يعد هذا شرطا من شروط تطبيق نموذج راش ، ويعتبر هؤلاء الطلبة ملائمين للنموذج ؛ أي أنهم ضمن المدى الذي يمكن لهذا الاختبار أن يميز بين مستوياتهم ، وعليه فإن جميع الطلبة البالغ عددهم (5520) قد دخلوا في عملية التحليل . كذلك لا توجد فقرة أجاب عنها جميع الطلبة ولا توجد أي فقرة أخفق جميع الطلبة في الإجابة عنها، وبالتالي اعتبرت جميع الفقرات العشرين فقرات ملائمة لاستجابة الطلبة وأدخلت جميعها في عملية التحليل .

ومـــن أجـــل الإجابـــة عـــن سؤال الدراسة الأول والمتعلق بمدى وجود علاقة ارتباطـــية ذات دلالة إحصائية بين تقديرات قدرات الأفراد المفحوصين حسب كل من الـــنظرية الكلاســـيكية ونمـــوذج راش ، قام الباحث بإيجاد العلامة الكلية لكل طالب " Total Score " والتـــي تمثل قدرته حسب النظرية الكلاسيكية ، وقيمة قدرته بوحدة اللوجيت حسب نموذج راش ، وقبل حساب معامل الارتباط باستخدام معادلة بيرسون تم التحقق من افتراض الخطية من خلال رسم شكل الانتشار لأنه عند حساب المعامل يفـــترض أن العلاقة بين المتغيرين تنزع نحو وجود علاقة خطية وليست إنحنائية لذا ينصح برسم شكل الانتشار خاصة في الحالات التي يصعب الحكم عليها بأن البيانات ترتبط بعلاقة خطية دون رسم شكل الانتشار(الخليلي و عودة ، 1988) .

تـــم إيجاد معامل ارتباط بيرسون بين العلامة الكلية لكل طالب وقيمة قدرته ، ثم تم تفسير قيمة هذا المعامل للتعرف على مدى قوة الارتباط واتجاه العلاقة.

أما الإجابة عن السؤال الثاني : والمتعلق بمدى وجود علاقة ارتباطية ذات دلالة إحصائية بين معاملات صعوبة الفقرات حسب كل من النظرية الكلاسيكية ونموذج راش ، تم التحقق أولا من افتراض الخطية من خلال رسم شكل الانتشار بين قيم صعوبة الفقرات وفق كل من النظرية الكلاسيكية ونموذج راش ، ثم تم إيجاد معامل ارتباط بيرسون بين هذه القيم .

أما الإجابة عن السؤال الثالث والمتعلق بمدى استقرار قيم قدرات الأفراد عند استخدام عينات مختلفة من الفقرات ، ومدى استقرار قيم معاملات الصعوبة للفقرات عند استخدام عينات مختلفة من الأفراد حسب كل من النظرية الكلاسيكية ونموذج راش فقد أشار رايت (Wright,2002) أنه للتحقق من استقرار قيم قدرات الأفراد المفحوصين وتحررها من أثر فقرات الاختبار " Item-Free " نستخدم مجموعتين من الفقرات : مجموعة من الفقرات ذات قيم صعوبة مرتفعة ومجموعة ذات قيم صعوبة منخفضة حيث تمثل كل مجموعة اختبارا فرعيا يتم تطبيقه على عينة واحدة من المفحوصين ، ثم نجد قدرات الأفراد المفحوصين من كلا الاختبارين ، ونجد العلاقة الارتباطية بين هذه القيم حسب كل من النظرية الكلاسيكية ونموذج راش ، ثم نرسم

شكل الانتشار بحيث إذا اقترب الرسم البياني للشكل من الخط المستقيم دل على ذلك تحرر قيم قدرات الأفراد من أثر فقرات الاختبار .

لذا تم تقسيم الاختبار مناصفة إلى اختبارين : الأول يتكون من (10) فقرات ذات قيم صعوبة مرتفعة ، والثاني يتكون من (10) فقرات ذات قيم صعوبة منخفضة ثم طبق كل اختبار على عينة الطلبة المفحوصين البالغ عددهم (5520) مفحوصا ، ثم تم إيجاد قدرة كل طالب وفق كل من النظرية الكلاسيكية ونموذج راش على الاختبارين(نصفي الاختبار) وحساب معامل الارتباط.

و للتحقق من استقرار قيم صعوبة الفقرات وتحررها من أثر عينة المفحوصين "person-free" ، يتم أخذ مجموعتين من المفحوصين يميزهما الاختبار : مجموعة ذات أداء مرتفع ومجموعة ذات أداء منخفض ويطبق عليهما الاختبار الذي يتكون من (20) فقرة ، ثم يتم حساب قيم صعوبة الفقرات من خلال كل مجموعة على حدة حسب كل من النظرية الكلاسيكية ونموذج راش ، وايجاد العلاقة الارتباطية بين هذه القيم ومن ثم نرسم شكل الانتشار بحيث إذا اقترب الرسم البياني للشكل من الخط المستقيم دل ذلك على تحرر قيم صعــوبة الفقرات من أثر عينة المفحوصين .

وعليه تم أخذ أداء أعلى (27%) وأداء أدنى (27%) من الطلبة المفحوصين على اعتبار أن هذه النسبة هي أفضل نسبة يميزها الأختبار كما أشار بذلك كيلي (Kelley,1939) ، حيث بلغ عدد أداء أفراد كل مجموعة (1490) طالبا ، ثم طبق الاختبار المكون من (20) فقرة على كل مجموعة على حدة ، و تم إيجاد صعوبة كل فقرة وفق النظرية الكلاسيكية ونموذج راش من خلال المجموعتين .

محددات الدراسة

هناك مجموعة من المحددات يجب أن تؤخذ أن تؤخذ بعين الاعتبار عند تعميم النتائج خارج عينة الدراسة وهي :

1 – اقتصـــار الدراســـة على فقرات اختبار الاختيار من متعدد الواردة في الاختبار الوطني .

الفصل الرابع

عرض النتائج

يتناول هذا الفصل عرضا لنتائج الدراسة والتي هدفت إلى مقارنة صعوبة الفقرة وقدرة الفرد في ضوء كل من النظرية الكلاسيكية ونموذج راش ، ويمكن عرض النتائج على النحو الآتي .

أولا : نتائج التحليل العاملي

كان الهدف من التحليل العاملي " Factor analysis "هو للتحقق من افتراض أحادية البعد للسمة المقيسة (التحصيل) ، وقد استخدمت طريقة تحليل المكونات الأساسية " Principal Component Analysis) " مع التدوير المتعامد " Varimax Rotation " .

ويبين الجدول رقم (3) نتائج التحليل العاملي الذي يظهر فيه قيم الجذر الكامن والتباين المفسر لكل عامل .

جدول رقم (3)

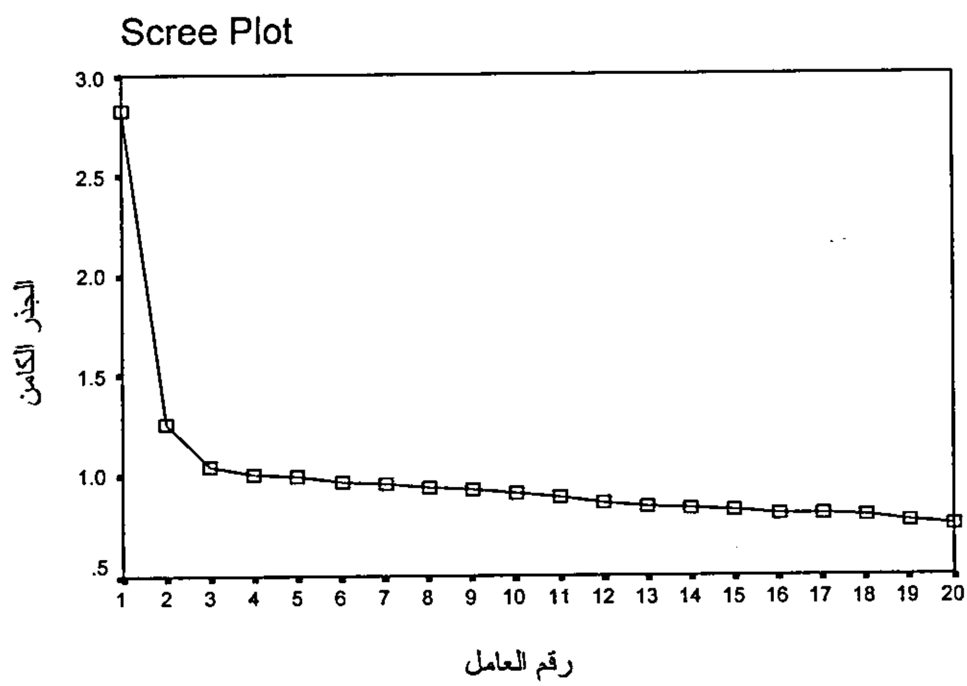
نتائج التحليل العاملي

العامل	الجذر الكامن	نسبة التباين المفسر	نسبة التباين المفسر الكلي
الأول	2.82	14.13	14.13
الثاني	1.15	6.26	20.39
الثالث	1.03	5.19	25.58
الرابع	1.00	5.02	30.60

ويتضح من هذا الجدول أن نسبة الجذر الكامن للعامل الأول إلى العامل الثاني تزيد عن (2) ، والجذر الكامن للعامل الثاني ليست أكبر بكثير من العاملين الثالث والرابع ، وأن العامل الأول يساهم في تفسير 14.13% من تباين درجات

المفحوصين ، وبناء عليه فإن فقرات الاختبار تحقق افتراض أحادية البعد للسمة المقيسة .

وباستخدام محك كاتل (Cattell) الذي يطلق عليه اختبار فرز العوامل" Scree Test " و يوضحه الشكل رقم (6) ، نلاحظ أن قيم الجذر الكامن تتناقص بشكل كبير بين العامل الأول والعامل الثاني إلى أن يصل إلى نقطة الانعطاف ثم يبدأ الجذر الكامن بالتناقص بصورة ضئيلة بحيث يقترب الخط البياني في الاستواء مع المحور الأفقي تقريبا ، مما يؤكد وجود عامل عام سائد وبالتالي تحقق فرض أحادية البعد لفقرات الاختبار محل الدراسة .



الشكل رقم (6)

اختبار فرز العوامل Scree Test

ثانيا : النتائج المتعلقة بالسؤال الأول

هل توجد علاقة ارتباطية ذات دلالة إحصائية بين تقديرات قدرات الأفراد المفحوصين حسب كل من النظرية الكلاسيكية ونموذج راش ؟

وللإجابة عن هذا السؤال يمكن ترتيب النتائج على النحو الآتي :

أ – تشير العلامة الكلية على فقرات الاختبار لكل مفحوص إلى قدرته وفق النظرية الكلاسيكية والجدول رقم (4) يبين عدد الطلبة المفحوصين ونسبتهم على العلامة الكلية للاختبار .

جدول رقم (4)

قيم قدرات الطلبة المفحوصين على الاختبار وفق النظرية الكلاسيكية

العلامة الكلية	عدد الطلبة	النسبة المئوية %
4	1	~ 0
5	2	~ 0
6	10	0.2
7	56	1.0
8	151	2.7
9	368	6.7
10	680	12.3
11	928	16.8
12	1059	19.2
13	985	17.8
14	637	11.5
15	389	7.0
16	176	3.2
17	59	1.1
18	15	0.3
19	4	0.1
المجموع	**5520**	**100**

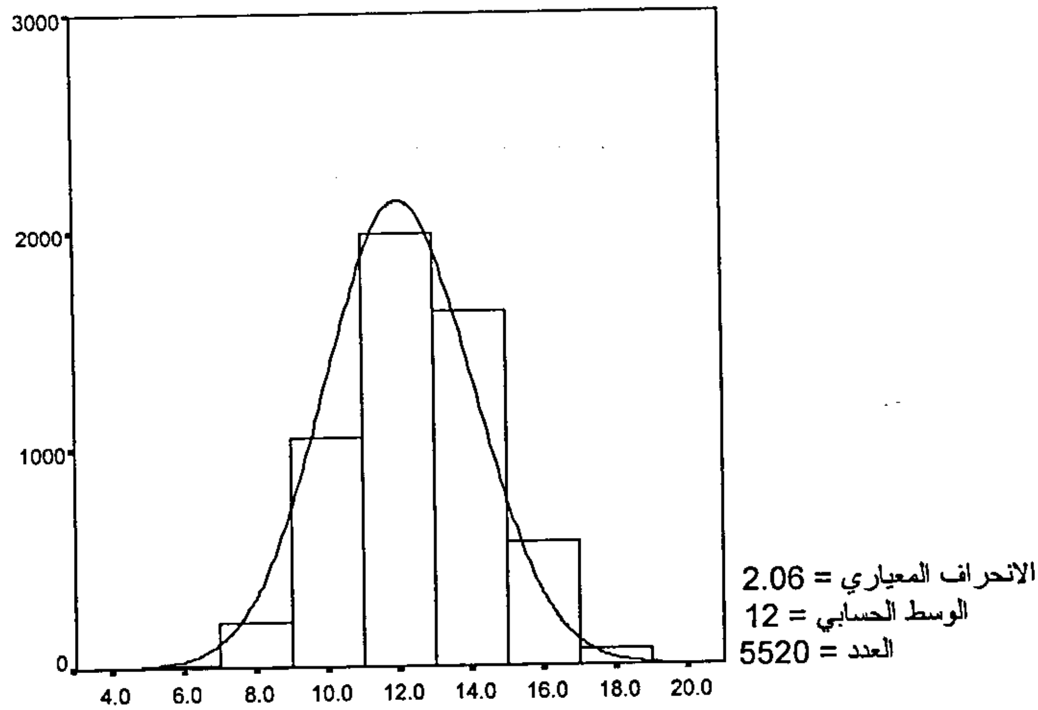
أما الجدول رقم (5) فيبين بعض قيم مقاييس النزعة المركزية والتشتت للعلامة الكلية على الاختبار .

جدول رقم (5)

قيم مقاييس النزعة المركزية والتشتت للعلامة الكلية على الاختبار

الوسط الحسابي	الانحراف المعياري	التباين	الوسيط	المنوال	الالتواء	المدى
12.03	2.06	4.24	12	12	0.02	15

ويتضح من نتائج الجدولين أن عدد الطلبة المفحوصين يساوي (5520)، تراوحت علاماتهم بين (4) الى (19) . ولا يوجد أحد من الطلبة حصل على العلامة النهائية (20) أو أخفق في الاختبار أي حصل على العلامة (صفر) ، ويلاحظ من الجدول (4) أن الوسيط يساوي المنوال ويساوي الوسط الحسابي تقريبا مما يشير إلى أن توزيع العلامات يقترب كثيرا من التوزيع الطبيعي كما يبينه الشكل رقم (7) .



شكل رقم (7)

توزيع علامات الطلبة على الاختبار

ب – تم تقدير قيم قدرات الطلبة المفحوصين وفق نموذج راش بوحدة تسمى اللوجيت (Logit) بقيم سالبة وموجبة كما هو موضح في جدول (6).

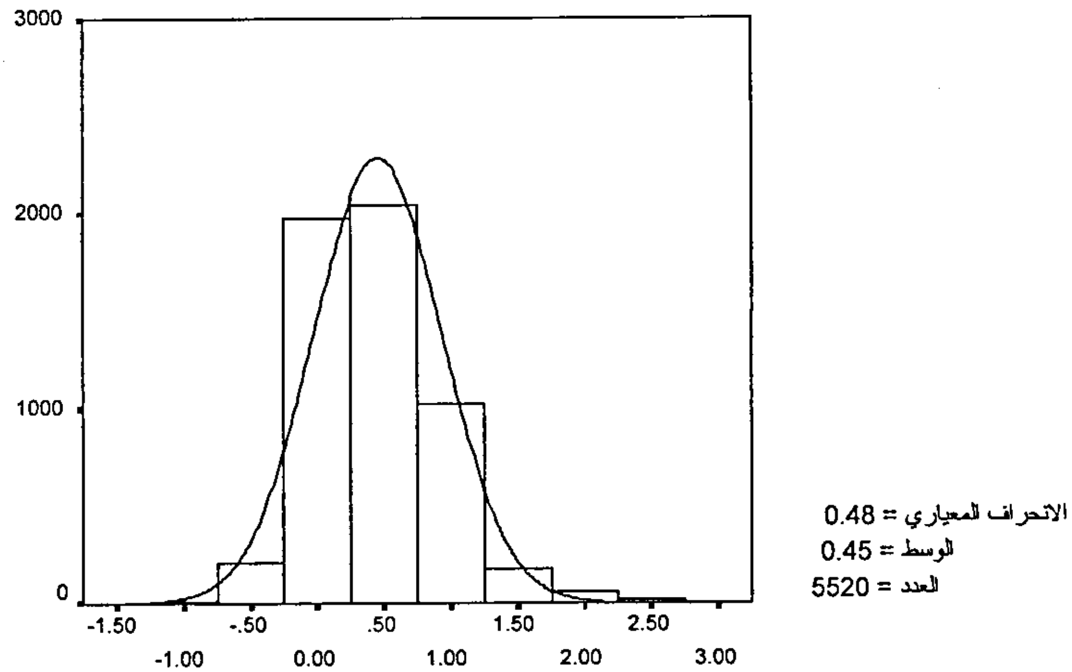
الجدول رقم (6)

قيم قدرات الطلبة المفحوصين وفق نموذج راش

القدرة (θ) بـــاللوجيت	عدد الطلبة	النسبة المئوية
-1.45	2	~0
-1.15	2	~0
-0.89	10	0.2
-0.65	56	1.0
-0.34	151	2.7
-0.21	368	6.7
0.0	680	12.3
0.21	927	16.8
0.42	1059	19.2
0.65	986	17.9
0.88	637	11.5
1.15	389	7.0
1.45	176	3.2
1.81	59	1.1
2.28	14	0.3
3.05	4	0.1
المجموع	**5520**	**100**

يتضح من هذا الجدول أن أقل قيمة للقدرة تساوي (-1.45) لوجيت وأعلى قيمة تساوي (3.05) لوجيت بوسط حسابي مقداره (0.45) لوجيت وانحراف معياري (0.48) لوجيت وقيم قدرات الطلبة تقع ضمن المدى (-3 و 3) لوجيت .

ويقترب توزيع هذه القيم من التوزيع الطبيعي الذي يبينه الشكل (8) .



شكل رقم (8)
توزيع قدرات الطلبة حسب نموذج راش

جـــ- لإيجاد معامل الارتباط بين قيم قدرات الطلبة المفحوصين تم مقابلة قيمة القدرة لكل طالب حسب النظرية الكلاسيكية مع قيمة قدرته حسب نموذج راش وذلك لجميع الطلبة البالغ عددهم (5520) والذي يوضحه الجدول رقم (7) استنادا للملحق (3) الخاص بملف مخرجات برنامج Winsteps المتعلق بقيم القدرة للطلبة المفحوصين .

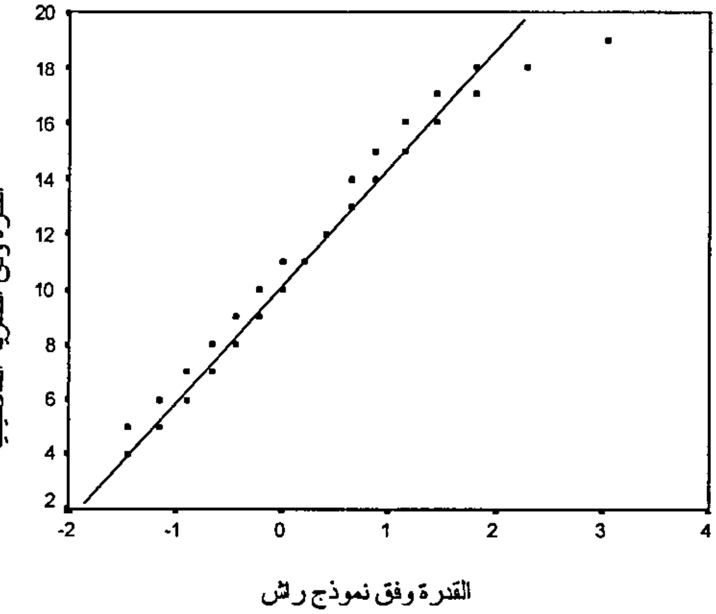
جدول (7)

قيم قدرات الطلبة وفق النظرية الكلاسيكية وما يقابلها في نموذج راش

وفق النظرية الكلاسيكية		وفق نموذج راش	
القدرة	العدد	القدرة	العدد
19	4	3.05	4
18	15	2.28	14
17	59	1.81	59
16	176	1.45	176
15	389	1.15	389
14	637	0.88	637
13	985	0.65	986
12	1059	0.42	1059
11	928	0.21	927
10	680	0.00	680
9	368	-0.21	368
8	151	-0.43	151
7	56	-0.65	56
6	10	-0.89	10
5	2	-1.15	2
4	1	-1.45	2
المجموع	5520	---------	5520

يتضح من الجدول (7) أنه كلما زادت قدرة المفحوص حسب النظرية الكلاسيكية زادت قدرته حسب نموذج راش والعكس صحيح أي أن العلاقة طردية ، فالطالب المفحوص الذي علامته على الاختبار تساوي (19) والتي تشير إلى قدرته وفق النظرية الكلاسيكية ، وجد أن قدرته تساوي (3.05) وحدة لوجيت وفق نموذج راش ، والطالب الذي حصل على العلامة (4) تقابلها قدرة تساوي (1.45-) لوجيت وهكذا . وكلما زادت سالبية القيمة حسب نموذج راش دل ذلك على نقصان في قدرة الطالب المفحوص ، فالطالب الذي قدرته (0.21-) لوجيت أعلى قدرة من الطالب الذي قدرته (1.45-) لوجيت في الإجـابة عن فقرات الاختبار .

ويتضح أيضا أن مدى تدرج قدرات الطلبة يمتد من (‎−1.45‎) إلى (3.05) لوجيت . وعند رسم شكل الانتشار بين قيم القدرة وفق النظرية الكلاسيكية ونموذج راش يتضح أن هذا الشكل ينزع إلى وجود علاقة خطية كما يبينها الشكل (9) .



شكل رقم (9)

شكل انتشار للعلاقة بين قيم القدرة

ثم تم حساب معامل ارتباط بيرسون بين علامة الطالب (قدرته) وفق النظرية الكلاسيكية وقدرته وفق نموذج راش ، وكانت النتيجة أن معامل الارتباط ذو دلالة إحصائية عند ($\alpha = 0.01$) ويساوي (0.994) . مما يدل على أن العلاقة قوية وطردية بين توزيع قيم القدرة التي تمثلها العلامة الكلية وفق النظرية الكلاسيكية ، وتوزيع قيم القدرة وفق نموذج راش .

ثالثا : النتائج المتعلقة بالسؤال الثاني :

هل توجد علاقة ارتباطية ذات دلالة إحصائية بين معاملات صعوبة الفقرات حسب كل من النظرية الكلاسيكية ونموذج راش ؟

وللإجابة عن هذا السؤال يمكن ترتيب النتائج على النحو الآتي :

أ – تم حساب معاملات الصعوبة لفقرات الاختبار والبالغ عددها (20) فقرة وفق النظرية الكلاسيكية وذلك بإيجاد نسبة الطلبة الذين أجابوا إجابة صحيحة عن كل فقرة .

ويبين الجدول (8) قيم معاملات الصعوبة للفقرات مرتبة ترتيبا تصاعديا .

جدول رقم (8)

معاملات صعوبة الفقرات وفق النظرية الكلاسيكية

رقم الفقرة	عدد الإجابة الصحيحة/ فقرة	معامل الصعوبة
9	1676	0.304
15	2671	0.484
10	2984	0.540
17	3005	0.544
5	3032	0.549
18	3061	0.555
6	3154	0.571
7	3238	0.586
12	3243	0.588
1	3289	0.596
2	3308	0.599
8	3329	0.603
3	3477	0.630
11	3511	0.636
20	3557	0.644
16	3600	0.652
19	3775	0.684
14	3972	0.719
13	4240	0.768
4	4308	0.780

يتضح من الجدول أن أصعب فقرة هي الفقرة رقم (9) حيث أن معامل الصعوبة لها يساوي (0.304) ، بالرغم أنها أقل معامل صعوبة بين الفقرات وهذا بسبب أن معامل الصعوبة في النظرية الكلاسيكية هو نسبة الذين أجابوا صحيحا على الفقرة .

وأسهل فقرة هي الفقرة رقم (4) والتي معامل صعوبتها يساوي (0.78) وتقع معاملات الصعوبة للفقرات بين (0.30) و(0.78) بوسط حسابي (0.60) أي ضمن معاملات الصعوبة المقبولة .

وإذا أخذ التخمين بعين الاعتبار للاختيار من أربعة بدائل ، فإن معامل الصعوبة المناسب يساوي تقريبا (0.63) ، وعليه فإن متوسط صعوبة فقرات الاختبار يقل بمقدار (0.03) مما يدل أن معظم فقرات الاختبار كانت ذات صعوبة متوسطة .

ب - لحساب معاملات الصعوبة وفق نموذج راش تم التأكد من أن الفقرات ملائمة لهذا النموذج وذلك من خلال الشكل (10) الذي يرمز لكل فقرة برمز النجمة (*) ، مع التوضيح بالرسم حدود الملائمة وموقع كل فقرة من هذه الحدود ، وذلك لمؤشري الملائمة : التقاربي (Infit) والتباعدي (Outfit) بحيث إذا وقعت الفقرة ضمن حدود الملائمة فان الفقرة يتم إدخالها في عملية التحليل ، ويلاحظ من خلال الشكل أن جميع الفقرات كانت ضمن حدود الملائمة وبالتالي أدخلت جميعها في التحليل .

```
|+---------------------------------------------------------------------------------+
ENTRY | MEASURE |   INFIT MEAN-SQUARE   | OUTFIT MEAN-SQUARE     |     |
|NUMBER| -      + |0     0.7 1 1.3     2|0     0.7 1 1.3      2| ITE|
|------+----------+-----------------------+-----------------------+----|
|     9|         *|        :  *  :        |        :  *  :        | q9 |
|    15|     *    |        :  *  :        |        :  *  :        | Q15|
|    10|    *     |        :  *  :        |        :  *  :        | q10|
|    17|    *     |        :  *  :        |        :  *  :        | Q17|
|     5|    *     |        :  *  :        |        :  *  :        | q5 |
|    18|    *     |        :  *  :        |        :  *  :        | Q18|
|     6|    *     |        :  *  :        |        :  *  :        | q6 |
|     7|   *      |        : *.  :        |        : *.  :        | q7 |
|    12|   *      |        :  *  :        |        :  *  :        | Q12|
|     1|   *      |        : *.  :        |        : *.  :        | q1 |
|     2|   *      |        : *.  :        |        : *.  :        | q2 |
|     8|   *      |        : *.  :        |        : *.  :        | q8 |
|     3|   *      |        : *.  :        |        : *.  :        | q3 |
|    11|   *      |        :  *  :        |        :  *  :        | Q11|
|    20|   *      |        : *.  :        |        : *.  :        | Q20|
|    16|  *       |        : *.  :        |        : *.  :        | Q16|
|    19|  *       |        : *.  :        |        : *.  :        | Q19|
|    14| *        |        : *.  :        |        : *.  :        | Q14|
|    13|*         |        : *.  :        |        : *.  :        | Q13|
|     4|*         |        : *.  :        |        : *.  :        | q4 |
+---------------------------------------------------------------------------------+
```

شكل رقم (10)

مدى ملائمة فقرات الاختبار لنموذج راش

ويبين الجدول (9) معامل صعوبة كل فقرة بوحدة اللوجيت (Logit)

الجدول رقم (9)

معاملات صعوبة الفقرات وفق نموذج راش باستخدام برنامج (winsteps)

رقم الفقرة	معامل الصعوبة	رقم الفقرة	معامل الصعوبة
9	1.32	2	0.03
15	0.52	8	0.01
10	0.28	3	-0.11
17	0.26	11	-0.14
5	0.24	20	-0.18
18	0.22	16	-0.21
6	0.15	19	-0.36
7	0.08	14	-0.54
12	0.08	13	-0.80
1	0.04	4	-0.88

ويتضح من الجدول أن معاملات الصعوبة تأخذ قيم سالبة وقيم موجبة ، وأشار كل من هامبلتون وسوامنثان (Hambeleton & Swamanthan, 1985) أن معاملات صعوبة الفقرة تتغير من (−2) إلى (+2) ؛ فالمعاملات التي تقترب من القيمة (−2) تعتبر فقرات سهلة جدا ، والتي تقترب من القيمة (+2) تعتبر صعبة جدا بالنسبة للطلبة المفحوصين . وعليه فإن أقل قيمة يظهرها الجدول هي (−0.88) وهي للفقرة رقم (4) و التي تعتبر أسهل فقرة ، وأكبر قيمة هي (1.32) ، وهي للفقرة رقم (9) و التي تعتبر أصعب فقرة .

ويتضح أيضا أن مدى تدرج الفقرات يقع ضمن المدى (+2 ، −2) بوسط حسابي يساوي (صفر) أي ضمن المدى المتوسط والمقبول لمعاملات الصعوبة .

جـ− لحساب معامل الارتباط بين معاملات الصعوبة لفقرات الاختبار تم مقابلة كل قيمة من قيم معامل الصعوبة حسب النظرية الكلاسيكية مع كل قيمة من قيم معامل الصعوبة حسب نموذج راش ولكل فقرة كما يوضحه الجدول (10) بعد عملية التقريب .

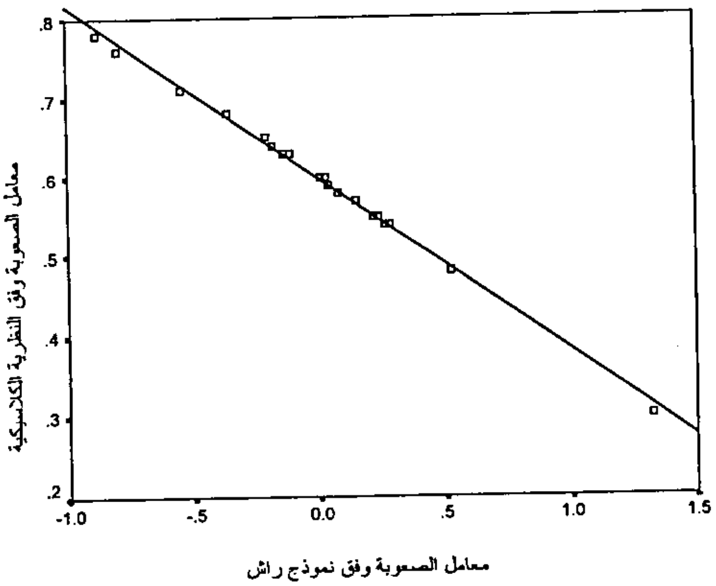
جدول (10)

معامل الصعوبة لكل فقرة وفق النظرية الكلاسيكية ونموذج راش

رقم الفقرة	معامل الصعوبة وفق النظرية الكلاسيكية	معامل الصعوبة وفق نموذج راش بوحدة "اللوجيت"
9	0.30	1.32
15	0.48	0.52
10	0.54	0.28
17	0.54	0.26
5	0.55	0.24
18	0.55	0.22
6	0.57	0.15
7	0.58	0.08
12	0.58	0.08
1	0.59	0.04
2	0.60	0.03
8	0.60	0.01
3	0.63	0.11-
11	0.63	0.14-
20	0.64	0.18-
16	0.65	0.21-
19	0.68	036-
14	0.71	0.54-
13	0.76	0.80-
4	0.78	0.88-

يتضح من الجدول السابق أن أصعب فقرة وفق كل من النظرية الكلاسيكية ونموذج راش كانت فقرة رقم (9) وأن أسهل فقرة كانت فقرة رقم (4) على الرغم من العلاقة العكسية بين هذه القيم ، وهذا بسبب المنطق المعكوس في تعريف معامل الصعوبة وفق النظرية الكلاسيكية ،

أما الشكل رقم (11) فيبين نوع العلاقة بين هذه القيم واتجاهها.



شكل رقم (11)

شكل انتشار يبين معاملات صعوبة الفقرة وفق نموذج راش

يظهر من هذا الشكل أن العلاقة تقريبا خطية وعكسية ، أما قوة هذه العلاقة فتتضح من خلال معامل الارتباط بين معاملات الصعوبة وفق كل من النظرية الكلاسيكية ونموذج راش ، فقد وجد أن قيمة هذا المعامل يساوي (0.998-) ، وهذا يشير الى ان العلاقة قوية جدا وشبه تامة ودالة إحصائيا عند ($\alpha = 0.01$) .

رابعا : النتائج المتعلقة بالسؤال الثالث :

هل قيم قدرات الأفراد مستقرة عند استخدام عينات مختلفة من الفقرات ؟ وهل قيم معاملات الصعوبة للفقرات مستقرة عند استخدام عينات مختلفة من الأفراد حسب كل من النظرية الكلاسيكية ونموذج راش ؟

و للتحقق من استقرار قيم قدرات الطلبة المفحوصين كانت النتائج كما يلي :

1 – قُسّم الاختبار الكلي المكون من (20) فقرة إلى اختبارين فرعيين ، أحدهما الاختبار السهل ويتكون من أسهل (10) فقرات من فقرات الاختبار الكلي ، والآخر الاختبار الصعب ويتكون من أصعب (10) فقــرات من فقرات الاختبار الكلي .

والجدول (11) يبين فقرات كل من الاختبارين السهل والصعب حسب معامل صعوبتها .

الجدول رقم (11)

معامل الصعوبة لفقرات نصفي الاختبار السهل والصعب

الاختبار السهل		الاختبار الصعب	
رقم الفقرة	معامل الصعوبة	رقم الفقرة	معامل الصعوبة
4	0.78	9	0.3
13	0.76	15	0.48
14	0.71	10	0.54
19	0.68	17	0.54
16	0.65	5	0.55
20	0.64	18	0.55
11	0.63	6	0.57
3	0.63	7	0.58
8	0.60	12	0.58
2	0.60	1	0.59

2 – حددت تقديرات القدرة للطلبة المفحوصين وفق النظرية الكلاسيكية بأن طبق الاختباران: السهل والصعب كل على حدة على عينة الطلبة البالغ عددهم (5520) ، وتم إيجاد العلامة الكلية على كل اختبار ولكل طالب مفحوص والتي تمثل قدرته على ذلك الاختبار وفق النظرية الكلاسيكية كما يظهرها الجدول (12) .

الجدول رقم (12)

قيم قدرات الطلبة المفحوصين على الاختبارين وفق النظرية الكلاسيكية

	العلامة	0	1	2	3	4	5	6	7	8	9	10
الاختبار السهل	عددالطلبة	0	4	28	142	363	770	1052	1282	1101	587	191
	نسبة منوية	0	0.1	0.5	2.6	6.6	13.9	19.1	23.2	19.9	10.6	2.5
الاختبار الصعب	عدد الطلبة	3	25	198	533	1006	1163	1245	860	389	95	6
	نسبة منوية	~0.0	0.5.	3.6	9.7	18.2	21.1	22.5	15.6	7.0	1.7	0.1

أما الجدول (13) فيبين مقارنة بين الاختبار السهل والاختبار الصعب من حيث قيم مقاييس النزعة المركزية والتشتت حسب النظرية الكلاسيكية .

الجدول رقم (13)

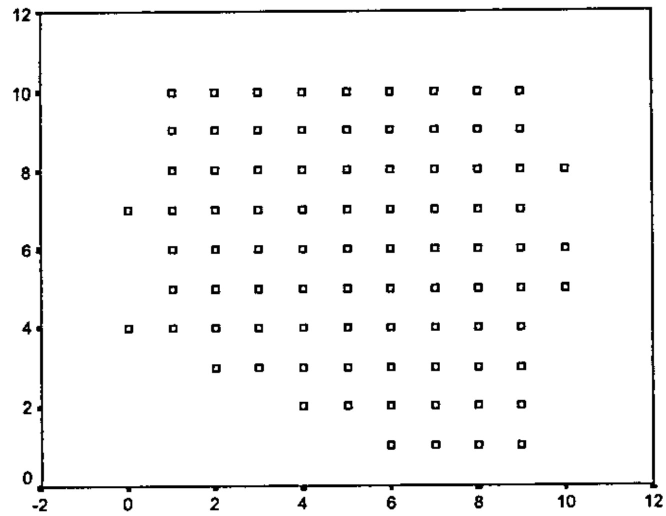
مقاييس النزعة المركزية والتشتت للاختبارين

	الوسط	الوسيط	المنوال	الانحراف المعياري	أدنى علامة	أعلى علامة
الاختبار السهل	6.7	7	7	1.66	1	10
الاختبار الصعب	5.3	5	6	1.63	صفر	10

يتضح من الجدولين أن متوسط علامات الاختبار السهل (6.7) أعلى من متوسط علامات الاختبار الصعب (5.3) ، وللتحقق من وجود فروق في المتوسطات بين أداء الطلبة على الاختبار السهل والاختبار الصعب ، تم إجراء اختبار (t- test) وأظهرت نتائج هذا الاختبار (الملحق رقم 4) وجود فروق ذات دلالة إحصائية عند (α=0.05)

كذلك حصل (191) طالبا على العلامة الكلية في الاختبار السهل بينما كان العدد في الاختبار الصعب وعلى العلامة نفسها هو (6) طلاب ، وهناك ثلاثة طلاب حصلوا على العلامة (صفر) في الاختبار الصعب بينما لا يوجد أي طالب حصل على هذه العلامة في الاختبار السهل .

وعند رسم شكل الانتشار بين علامات الطلبة على الاختبارين أظهر الشكل صعوبة في تحديد أي اتجاه للعلاقة بين العلامات لأن النقاط تشكل شكلا شبه دائري كما يظهره الشكل (12) .



شكل رقم (12)

شكل انتشار يبين قدرات الطلبة على الاختبارين السهل والصعب

وعند حساب معامل الارتباط بين علامات الطلبة على الاختبارين – أي علامة الطالب على الاختبار السهل وعلامتة على الاختبار الصعب – تبين أن قيمة معامل الارتباط تساوي (−0.22) وله دلالة إحصائية عند ($\alpha = 0.01$) .

ويشير ذلك إلى أن العلاقة بين علامات الطلبة على الاختبارين ضعيفة بالرغم من وجود دلالة إحصائية وذلك لأن حجم العينة كبير. وأن هناك عدم استقرار في قدرة كل طالب على الاختبارين ن ، وأن إشارة معامل الارتباط السالبة تعني أن علامات الغالبية العظمى من الطلبة كانت على الاختبار السهل أعلى منها على الاختبار الصعب .

3 – أما نتائج الطلبة المفحوصين على كلا الاختبارين وفق نموذج راش فقد تبين من خلال تقديرات القدرة للطلبة على الاختبار السهل ، أن (191) طالبا قد حصلوا على العلامة النهائية (10) ، وبالتالي لم يدخلوا في عملية التحليل ، ولم يوجد أي طالب حصل على العلامة (صفر) على الاختبار السهل .

وعليه فإن (5329) طالبا من أصل (5520) تم التعامل معهم من خلال الرزمة الإحصائية " winstep "من أجل إيجاد قيم قدراتهم .

ويبين الملحق رقم (5) جزء من النتائج العليا والدنيا لقيم القدرة على الاختبار السهل ، حيث يظهر في أعلى الجدول عدد الطلبة الذين تم إدخالهم وعدد الطلبة

الذين تم تحليلهم وعدد الفقرات التي أجابوا عنها بحيث أن أي طالب حصل على العلامة النهائية أشير إليه بعبارة "Maximum Estimated Measure " واستثني من التحليل ، ولم يوجد أي طالب حصل على العلامة (صفر) على الاختبار السهل .

أما تقديرات القدرة على الاختبار الصعب للطلبة المفحوصين فقد تبين (الملحق رقم 6) أن تسعة من الطلبة قد استثنوا من التحليل وذلك لأن ستة منهم حصلوا على العلامة النهائية " Maximum Estimated Measure "وثلاثة حصلوا على العلامة (صفر) " Minimum Estimated Measure "، وعليه يكون عدد الطلبة الذين أدخلوا في عملية التحليل (5511) طالبا .

والجدول (14) يبين تكافؤ العلامة الكلية والقدرة للمفحوصين على الاختبارين السهل والصعب على التوالي .

الجدول رقم (14)

قيم القدرة وما يكافؤها من العلامات على الاختبارين السهل والصعب وفق نموذج راش

العلامة	الاختبار السهل		الاختبار الصعب	
	القدرة	الخطأ المعياري	القدرة	الخطأ المعياري
0	3.53-*	1.86	3.53-*	1.86
1	2.24-	1.03	2.25-	1.06
2	1.42-	0.80	1.43-	0.80
3	0.87-	0.70	0.88-	0.70
4	0.41-	0.65	0.43-	0.65
5	0.00	0.64	0.01-	0.64
6	0.42	0.65	0.41	0.66
7	0.87	0.70	0.87	0.70
8	1.42	0.80	1.43	0.80
9	2.24	1.06	2.26	1.07
10	3.52*	1.86	3.56*	1.86

* قيم القدرة التي تقع خارج المدى [+3 ، -3]

يتضح من هذا الجدول ما يأتي :

1 – تمتد العلامة الكلية لأفراد العينة على الاختبار السهل من العلامة (1) إلى العلامة (10) وعلى الاختبار الصعب من العلامة (صفر) إلى العلامة (10) .

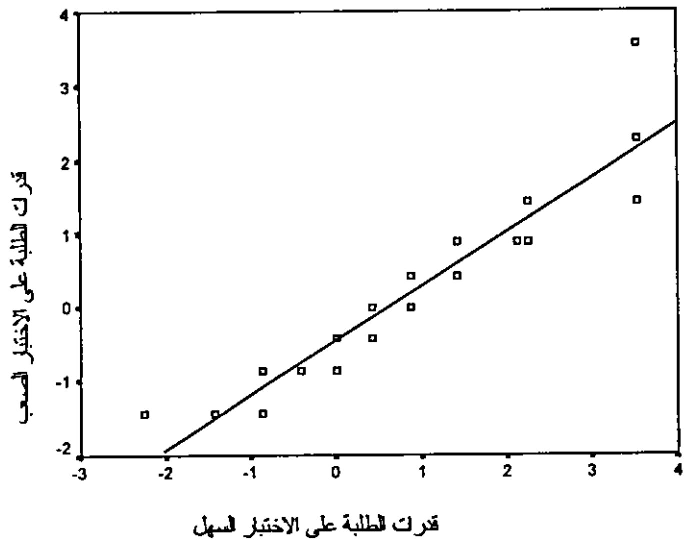
2 – يمتد مدى القدرة على الاختبار السهل من القيمة (–3.53) لوجيت إلى (3.52) لوجيت . و على الاختبار الصعب من القيمة (–3.53) لوجيت إلى (3.56) لوجيت .

3 – تقابل كل علامة خام ما يكافؤها من قيمة القدرة ولكل اختبار على حدة .

ويتضح أيضا أن هناك تكافؤا في قيم القدرة بين الاختبارين بغض النظر عن مستوى الاختبار سواء أكان سهلا أم صعبا،فالطالب الذي علامته الخام تساوي (6) كانت قيمة قدرته على الاختبارين : السهل والصعب تساوي (0.42) لوجيت و (0.41) لوجيت على التوالي وهكذا ، بمعنى أن قيم القدرة تقترب كثيرا من بعضها البعض .

وللتحقق من وجود فروق في المتوسطات بين أداء الطلبة على الاختبارين وفق نموذج راش ، فقد تم إجراء اختبار (t-test) عند دلالة إحصائية ($\alpha=0.05$) وأظهرت النتائج (ملحق7) إلى وجود فروق في المتوسطات .

وعند رسم شكل الانتشار كما في الشكل (13) بين قدرات الطلبة على الاختبارين السهل والصعب ، أظهر الشكل وجود علاقة خطية قوية .



شكل رقم (13)

شكل انتشار بين قدرات الطلبة على الاختبارين وفق نموذج راش

أما معامل الارتباط بين هذه القيم ومن خلال أفراد العينة فيساوي (0.90) وهي علاقة قوية جدا ودالة إحصائيا عند ($\alpha = 0.01$) وهذا يدل على أن هناك استقرارا إلى حد ما بين قيم قدرات الطلبة مما يشير إلى تحرر قدرات عينة المفحوصين من مجموعة الفقرات التي تم اختبارهم عليها .

و للتحقق من استقرار قيم معاملات صعوبة الفقرات ، تم أخذ أداء أعلى (27%) وأداء أدنى (27%) ، من عينة المفحوصين ، لأن هذه النسبة هي أفضل نسبة يميزها الاختبار كما اشار كيلي(kelley,1939) في دراسته ، حيث بلغ عدد أفراد كل عينة فرعية (1490) طالبا . وتم اختبار العينتين الفرعيتين على فقرات الاختبار الكلي البالغ عددها (20) فقرة ، وكانت النتائج على النحو الذي يظهره الملحق رقم (8) الناتج عن تحليل استجابات المفحوصين باستخدام الرزمة الإحصائية (winstep) .

والجدول (15) يبين معاملات صعوبة الفقرات وفق النظرية الكلاسيكية ونموذج راش وذلك لكل فقرة .

الجدول رقم (15)

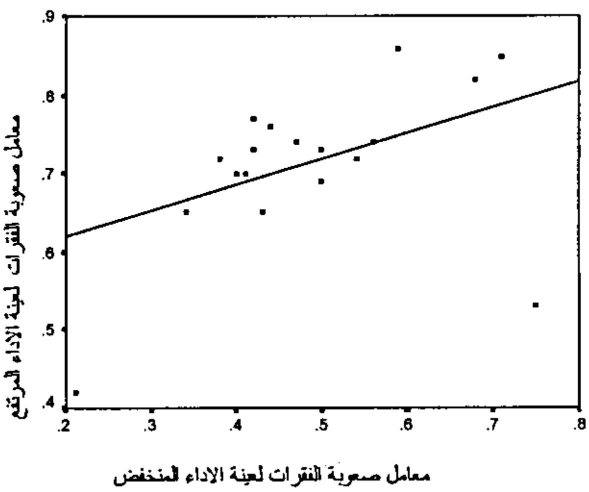
معاملات صعوبة الفقرات من خلال أداء العينتين المرتفع والمنخفض وفق النظرية الكلاسيكية ونموذج راش

رقم الفقرة	النظرية الكلاسيكية		نموذج راش	
	الأداء المرتفع	الأداء المنخفض	الأداء المرتفع	الأداء المنخفض
9	0.42	0.21	1.38	1.25
5	0.65	0.43	0.40	0.20
15	0.65	0.34	0.40	0.56
2	0.69	0.50	0.21	0.08
18	0.70	0.41	0.17	0.26
17	0.70	0.40	0.15	0.30
11	0.72	0.54	0.08	0.27-
10	0.72	0.38	0.07	0.38
6	0.73	0.50	0.02	0.14-
3	0.73	0.42	0.01-	0.23
1	0.74	0.47	0.01-	0.04
16	0.74	0.56	0.06-	0.34-
20	0.53	0.75	0.07-	0.21-
12	0.76	0.44	0.13-	0.13
8	0.76	0.44	0.15-	0.14
7	0.77	0.42	0.19-	0.23
19	0.70	0.40	0.25-	0.38-
13	0.82	0.68	0.52-	0.86-
4	0.85	0.71	0.71-	1.00-
14	0.86	0.59	0.77-	0.46-

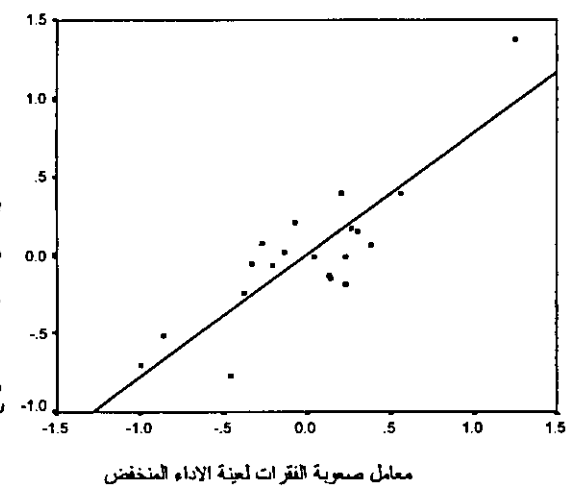
وعند حساب معامل الارتباط بين معاملات الصعوبة للفقرات الناتجة عن أداء العينتين : المرتفع والمنخفض حسب النظرية الكلاسيكية وجد أنه يساوي (0.43) ولم تكن له دلالة إحصائية ($\alpha = 0.01$) ، أما معامل الارتباط بين معاملات الصعوبة وفق نموذج راش فوجد أنه يساوي (0.86) وله دلالة إحصائية عند ($\alpha = 0.01$) وعند إجراء اختبار ت (T-test) – الملحق 9 – لمعرفة دلالة الفروق في المتوسطات لمعاملات الصعوبة للفقرات الناتجة عن أداء العينتين وجد أن هناك

فروقا ذات دلالة إحصائية ($\alpha = 0.01$) وفق النظرية الكلاسيكية بين المتوسطات ، أما وفق نموذج راش فلا توجد فروق ذات دلالة إحصائية.

ويظهر الشكلان (14أ ، 14ب) معاملات صعوبة الفقرات لكل من أداء العينتين المرتفع والمنخفض وفق كل من النظرية الكلاسيكية ونموذج راش .



شكل رقم(14أ)
وفق النظرية الكلاسيكية



شكل رقم (14ب)
وفق نموذج راش

يتضح من خلال شكلي الانتشار وقيمة معاملي الارتباط أن العلاقة بين معاملات صعوبة الفقرات الناتجة عن أداء العينتين كانت وفق نموذج راش تنزع نحو الخط المستقيم بصورة أوضح من النظرية الكلاسيكية ويؤكد ذلك قوة معامل الارتباط الذي يساوي (0.86) ذي الدلالة الإحصائية عند ($\alpha = 0.01$) ، وعليه يمكن الاستنتاج أن هناك استقرارا إلى حد ما بين معاملات صعوبة الفقرات وفق نموذج راش مما يشير إلى تحرر قيم معاملات صعوبة الفقرات من أفراد العينة الذين تم اختبارهم بها .

الفصل الخامس

مناقشة النتائج والتوصيات

المناقشة

هدفت هذه الدراسة إلى مقارنة صعوبة الفقرة وقدرة الفرد في ضوء كل من النظرية الكلاسيكية ونموذج راش باستخدام فقرات الاختيار من متعدد للاختبار الوطني للصف التاسع الأساسي في مبحث العلوم ، وحددت مشكلة الدراسة بثلاثة أسئلة رئيسية ، تم صوغ السؤال الأول على النحو التالي :

هل توجد علاقة ارتباطية ذات دلالة إحصائية بين تقديرات قدرات الأفراد حسب كل من النظرية الكلاسيكية ونظرية استجابة الفقرة من خلال نموذج راش .

أظهرت النتائج باستخدام البرمجيتين "SPSS" و"Winstep" أن قيم قدرات الطلبة على الاختبار وفق النظرية الكلاسيكية (جدول 3) أخذت قيما ليس فيها كسورا عشرية أو قيما سالبة وتراوحت بين (4) و(19) علامة وهذا ناتج عن الطريقة التي تم احتساب مؤشر القدرة "العلامة الكلية" بها .

أما في نموذج راش فإن القدرة للطلبة المفحوصين أخذت قيما سالبة وأخرى موجبة تحوي كسورا عشرية ولها وحدة قياس تسمى لوجيت "Logit" وتراوحت بين (‎-1.45‎) و(‎+3.05‎) لوجيت ، وهذا بسبب استخدام نموذج راش للدالة اللوغاريتمية "Logarithim Function" في حساب القدرة.

ومن المعلوم أن الطالب المفحوص يأتي إلى قاعة الاختبار ومعه قدرته على ذلك الاختبار ، والذي يهدف إلى تحديد موضعه على متصل القدرة الذي يقيسه الاختبار، يلاحظ من خلال نتائج قيم القدرة وطريقة احتسابها في النظرية الكلاسيكية أن علامات الطالب المفحوص والتي تعبر عن قدرته مقيدة بين قيمتين : العلامة النهائية الكاملة والعلامة الصفر ، وهاتان القيمتان تعبّران وفق هذه النظرية على أنها علامات تمثل القدرة التامة أو انعدامها.

بينما القيمتان (العلامة التامة والصفر) في نموذج راش تميزان على انهما تحويان معلومات قليلة عن موضع الطالب المفحوص على متصل السمة على اعتبار أن الطالب الذي يحصل على علامة كاملة أو يحصل على صفر الاختبار تكون قدرته أعلى أو أدنى من مستوى الاختبار الذي يقيس تلك القدرة فلا نستطيع معرفة وضعه على متصل السمة ، وبالتالي لا تتحدد قيم القدرة بقيمة نهائية دنيا أو قيمة نهائية عليا وانما تمتد عبر قيم لانهائية من – ∞ إلى + ∞ نظريا ، وعمليا تمتد من + 3 إلى – 3 مما يعكس قدرة المفحوص الواقعية .

وأظهرت النتائج أيضا – من خلال المقارنة بين قيم القدرة – اتفاقا عاليا في توزيع القدرة للطلبة المفحوصين بين النظرية الكلاسيكية ونموذج راش وذلك من خلال معامل الارتباط بين تقديرات القدرة التي تساوي (0.99) .

وهذا يعني أن توزيع القدرة للمفحوصين على درجة عالية من التشابه بين النظرية الكلاسيكية ونموذج راش ، وأن هذا الارتباط العالي يشير إلى أن القرارات حول مستويات المفحوصين تتشابه كثيرا بغض النظر عن الإطار النظري الذي أخذ منه القياس سواء النظرية الكلاسيكية أم نموذج راش ، علما أن توزيع تقديرات القدرة وفق نموذج راش لا تتطلب التوزيع الطبيعي "Normal Distribution" كما في النظرية الكلاسيكية. ويرى الباحث أن هذه النتيجة تشير إلى صحة الافتراضات التي انطلق منها نموذج راش ؛ بمعنى لو أن الارتباط كان ضعيفا بين قيم القدرة وفق النظرية الكلاسيكية ونموذج راش لقلنا أن منطلقات هذا النموذج غير واضحة المعالم أو أنها تتحدث عن موضوع آخر.

وهذه النتيجة تتفق مع نتائج دراسة كل من ماكدونالد وبوانن (Macdonald & Paunone, 2002) حيث كان معامل الارتباط بين قيم قدرات الأفراد في هذه الدراسة تساوي (0.97) ، وتتفق أيضا مع دراسة جمحاوي (2000) ، ومع نتائج دراسة تروسكوسكي (Truskosky, 1999) ، ودراسة أندرسون (Anderson, 1998) ، ودراسة فان (Fan, 1998) .

وقد تم صوغ السؤال الثاني في هذه الدراسة على النحو الآتي :

هل توجد علاقة ارتباطية ذات دلالة إحصائية بين معاملات صعوبة الفقرات حسب كل من النظرية الكلاسيكية ونظرية استجابة الفقرة من خلال نموذج راش ؟

للإجابة عن هذا السؤال تم أيضا استخدام البرمجيتين "SPSS" و"Winstep" حيث تم الحصول على معاملات صعوبة الفقرات وفق النظرية الكلاسيكية عن طريق إيجاد نسبة الذين أجابوا عن الفقرة إجابة صحيحة .

أظهرت النتائج أن معاملات صعوبة الفقرات وفق النظرية الكلاسيكية (جدول 11) تراوحت بين (0.30) و(0.78) للفقرتين (9) و(4) على التوالي ، ولا تحتوي المعاملات على قيم سالبة ، وبالتالي فإن حدود هذه القيم تقع بين (الصفر) و(الواحد صحيح) على أساس لو أن جميع الطلبة أجابوا إجابة صحيحة عن فقرة ما فإن معامل صعوبتها يساوي واحد ، ولو أن جميعهم أجابوا إجابة خاطئة عن فقرة ما فإن معامل صعوبتها يساوي صفر .

أما في نموذج راش فإن معاملات الصعوبة أخذت قيما سالبة وأخرى موجبة وأنها وقعت في مدى يتجاوز الواحد الصحيح ، والصفر فكانت أقل قيمة لمعامل الصعوبة هي للفقرة رقم (4) وتساوي (−0.88) وأكبر قيمة للفقرة (9) وتساوي (1.32) لوجيت .

إن معامل صعوبة الفقرة وفق نموذج راش غير مقيدة بحدود وانما تأخذ قيم من −∞ إلى+∞ نظريا، وعمليا من +2 إلى − 2 الجدول (11) .

وتمتاز قيم معاملات الصعوبة وفق نموذج راش بأن لها نفس وحدة القدرة وهي اللوجيت "Logit" وذلك لأن صعوبة الفقرة تقع على نفس تدريج مقياس القدرة وبالتالي فإن طريقة إيجاد معامل الصعوبة هي باستخدام الدالة اللوغاريتمية "Logarithim Function" .

ويعتبر معامل الصعوبة (الصفر أو الواحد) في النظرية الكلاسيكية أن الفقرة أما صعبة تامة الصعوبة أو سهلة تامة السهولة على التوالي وهذا النوع من الفقرات يتم التعامل معها ، بينما وفق نموذج راش تعد الفقرة التي يخفق في الإجابة عنها

جميع الطلبة المفحوصين بأن مستواها أعلى من مستوى قدراتهم ، والتي يجيب عنها جميع الطلبة المفحوصين فيعتبر مستواها أدنى من مستوى قدراتهم وهذا النوع من الفقرات يعتبر غير ملائمة لنموذج راش فلا يتم التعامل معها بل يتم حذفها.

وأظهرت النتائج وجود ارتباط عال بين معاملات الصعوبة وفق كل من النظرية الكلاسيكية ونموذج راش والذي يساوي (0.99−) عند دلالة إحصائية ($\alpha = 0.01$) ، أي أن توزيع معاملات الصعوبة وفق النظرية الكلاسيكية تتفق مع توزيع معاملات الصعوبة وفق نموذج راش . وهذه النتيجة تعزز صحة افتراضات النظرية الكلاسيكية ومصداقية نموذج راش في انتقائهما لفقرات الاختبار.

وهذه النتيجة تتفق مع دراسة هوانج (Hwang, 2002) ، ودراسة جمحاوي (2000) ، ودراسة ستيج (Stage, 1999) ، ودراسة فان (Fan, 1998) ، ودراسة هاشوي (Hashway, 1977) .

أما السؤال الثالث من أسئلة الدراسة فقد تم صوغه على النحو التالي :

هل قيم قدرات الأفراد مستقرة عن استخدام عينات مختلفة من الفقرات ؟ وهل قيم معاملات الصعوبة للفقرات مستقرة عند استخدام عينات مختلفة من الأفراد حسب كل من النظرية الكلاسيكية ونموذج راش ؟

أظهرت النتائج أن معامل الارتباط بين قيم قدرات الأفراد وفق النظرية الكلاسيكية باستخدام اختبارين فرعيين أحدهما سهل والآخر صعب يساوي (0.22−) وذو دلالة إحصائية عند ($\alpha = 0.01$) ، بينما معامل الارتباط بين قيم قدرات الأفراد على الاختبارين وفق نموذج راش يساوي (0.90) وذو دلالة إحصائية عند ($\alpha = 0.01$)، وأن العلاقة وفق نموذج راش بيـــن قيم القدرة تنزع نحو الخط المستقيم .

كذلك أظهرت النتائج أن معامل الارتباط بين قيم معاملات صعوبة الفقرات عند استخدام عينتين من الأفراد إحداهما ذات أداء عال والأخرى ذات أداء منخفض وفق النظرية الكلاسيكية يساوي (0.43) وليست لها دلالة إحصائية عند ($\alpha = 0.01$) ، أما معامل الارتباط وفق نموذج راش يساوي (0.86) وله دلالة

إحصائية عن ($\alpha = 0.01$) ، وأن العلاقة بين معاملات الصعوبة من خلال العينتين تنزع نحو وجود علاقة طردية .

ومن خلال قيم معاملات الارتباط بين قدرات الطلبة المفحوصين باستخدام اختبارين فرعيين أظهرت النتائج أن هناك استقرارا بين قيم قدرات الطلبة المفحوصين وفق نموذج راش أعلى منه وفق النظرية الكلاسيكية ، و أظهرت النتائج أيضا أن معاملات صعوبة الفقرات من خلال عينتين إحداهما ذات أداء مرتفع والأخرى ذات أداء منخفض كانت ذات استقرار أعلى وفق نموذج راش من النظرية الكلاسيكية.

وعليه يمكن القول أن نموذج راش يمتاز بأنه يحرر قيم قدرات الأفراد من أثر الفقرات ويحرر قيم معاملات صعوبة الفقرات من أثر عينة المفحوصين وهذه الميزة غير متوافرة في النظرية الكلاسيكية ، مما يجعل القياس من خلال نموذج راش وبالتالي من خلال نظرية استجابة الفقرة أقرب إلى الموضوعية منه من خلال النظرية الكلاسيكية ، وهذه النتيجة تتفق مع نتائج دراسة كل من ماكدونالد وبوانن (Macdonald & Paunone, 2002)، ودراسة فان (Fan, 1998) ، ودراسة ستيج . (Stage, 1999)

مما سبق و بالرغم من التشابه الكبير في النتائج من خلال مناقشة السؤالين الأول والثاني ، إلا أن ما يميز نموذج راش عن النظرية الكلاسيكية هو استقلالية قدرة الفرد المفحوص عن الفقرات التي يختبر بها "Item – Free" ، واستقلالية الفقرات عن عينة الأفراد المفحوصين "Person –Free" كما تبين في مناقشة السؤال الثالث ، بينما في النظرية الكلاسيكية لا تتوفر هذه الاستقلالية بل تعتمد قدرات الأفراد على عينة الفقرات وتعتمد صعوبة الفقرات على عينة الأفراد ، وبالتالي نحتاج إلى مزيد من فحص أوجه التشابه والاختلاف بين النظرية الكلاسيكية ونظرية استجابة الفقرة -الذي كان نموذج راش أحد نماذجها- وخاصة في خاصية عدم الاستقلالية بين قدرات الأفراد وصعوبة الفقرات وعدم تحققها في النظرية الكلاسيكية.

كذلك ما يميز نموذج راش هو انه يضع تقديرات القدرة وصعوبة الفقرات على مقياس مشترك وبوحدة القياس نفسها وهي اللوجيت "Logit" وذلك كما ظهر في نتائج الدراسة باستخدام برنامج "Winsteps" ، وهذه لها فائدة من ناحيتين :

الأولى: يمكننا التنبؤ بمدى قدرة الفرد على الإجابة عن فقرة ما بحيث إذا كانت قيمة قدرته المقدرة بوحدة اللوجيت اكبر من صعوبة الفقرة والمقدرة أيضا بوحدة اللوجيت فهناك احتمال كبير أن يجيب عن تلك الفقرة .

الثانية: يمكننا من تكوين بنك للأسئلة الذي يضم عدة اختبارات، تتدرج فقراتها جميعا في تدريج واحد مشترك .

وبالرغم من هذه المزايا لنموذج راش ، إلا انه ليس من السهولة انتقاء فقرات تتلاءم مع هذا النموذج فهو متحفظ جدا في اختياره للفقرات ، وخاصة في افتراضه أن جميع الفقرات يجب أن تتساوى في قدرتها على التمييز.

التوصيات:

بناء على نتائج الدراسة فانه يمكن وضع التوصيات الآتية :

1 – الاستفادة من كلا الطريقتين – النظرية الكلاسيكية ونموذج راش– في انتقاء الفقرات عند بناء الاختبار بحيث يبدأ اولا باستخدام النظرية الكلاسيكية لأن افتراضاتها بسيطة واي بيانات يمكن ان تحققها ، ثم باستخدام نموذج راش يتم اعادة انتقاء هذه الفقرات بحيث نصل الى فقرات ذات دقة عالية في خصائصها .

2 – هذه الدراسة استندت الى فقرات اختبار بني على أساس النظرية الكلاسيكية ، لذا ببناء اختبار تحصيلي وفق نموذج راش ثم عمل مقارنة لخصائص الفقرات وقدرات الافراد مع النظرية الكلاسيكية .

المراجع

أ- المراجع العربية

ثورندايــك ، وهيجــن. (1986) . القياس والتقويم في التربية وعلم النفس . ترجمة عبدالله زيد الكيلاني ، وعبدالرحمن عدس . مركز الكتب الأردني، عمان .

جمحاوي ، ايناس محمود .(200). مقارنة خصائص الفقرات وفق النظرية التقليدية ونظــرية اســتجابة الفقــرة فــي مقــياس للقــدرة الرياضــية . رســالة ماجستيرغيرمنشورة .جامعةاليرموك ، اربد

عــلاّم ، صـــلاح . (1995) . الأختبارات التشخيصية مرجعية المحك . دار الفكر العربي ، القاهرة .

عــودة ، احمــد ســليمان . (1992) . مــدى التوافق بين نموذج راش والمؤشرات التقليدية في اختيار مقياس اتجاه سباعي التدريج . حولية كلية التربية ، جامعة الامارات، العدد (8) ، ص (153 – 179) .

عــودة ، احمد سليمان . (1998) . القياس والتقويم في العملية التدريسية . ط2 ، دار الأمل ، اربد .

عــودة ، احمد سليمان ؛ والخليلي ، خليل يوسف . (1988) . الاحصاء للباحث في التربــية والعلــوم الانســانية ، (ط1) . دار الفكر للنشر والتوزيع ، عمان ، الاردن .

كــاظم ، أميــنة محمــد . (1986) . دراسة نقدية حول القياس الموضوعي للسلوك " نموذج راش" ، مؤسسة الكويت للتقدم العلمي ، الكويت .

ب - المراجع الأجنبية

Ackerman, T.A.(1999). Practical application of item response theory. Sponsered by the Department of Measurement Statistics and Evaluation Seminar and Training Center at the University of Maryland at College Park.

Allen, M.J. & Yen, W.M. (1979). Introduction to measurement theory. Monterey, California: Brooks/Cole.

Anastasi, A. (1988). Pyschological Testing. MacMillan Publishing: New-York.

Anderson,O.(1998). Does complex analysis (IRT) pay any dividends in achievement testing?. Paper Presented at Measurement and Evaluation, Conference Organized by Center for Research, University of Alberta.

Baker, F. B.(1985). The basics of item response theory. Portsmouth. NH : Heinmann.

Crislip, a., & Chin-Chance, Selvin. (2001). Using traditional psychometric methodogies and the rasch model in designing atest . Paper Presented at the Annual Meeting of the American Educational Research Assosiation (Seattle,WA, April 10-14,2001).

Crocker, L.M. & Algina, J. (1986). Introduction to classical and modern test theory. New York: Holt, Rinehart, and Winston.

Douglass and Frazier, M.(1979). Acomparison of classical and latent trait item analysis procedure. Educational and Psychological Measurement,Vol. 39,PP (33-352).

Ebel, R .(1979). Essentials of educational Measurement, (3th ed.) N.J:Engleswood Cliffs.

Fan, X. (1998). Item response theory and classical test theory: An empirical comparison of their item/person statistics. Educational and Psychological Measurement,Vol..(58) , No. (3), PP(357-381).

Fisher,Gerhard H. (1995), The derivation of Polytomous Rasch Model, in rasch models, Foundations, Recent Developments, and Applications, G.H.Fisher ,eds. PP(293-306) NewYork: Springer.

Hambleton, R.K. & Jones, R.W. (1993). Comparison of classical test theory and item response theory and their applications to test development. Educational Measurement: Issues and Practice, Vol.12, No(3),PP (38-47).

Hambleton, R.K., & Swaminathan, H. (1985). Item response theory: Principles and applications. Boston: Kluwer.

Hambleton, R.K., Swaminathan, H., & Rogers, H.J. (1991). Fundamentals of item response theory. Newbury Park, California: Sage Publications.

Hashway, R.M.(1977). Acomparison of tests derived using rasch and traditional psychometric paradigms. (doctoral dissertation, Boston College). Dessertation Abstract International, Vol. 38, No (2), PP (744-A).

Hulin, C .L.,Lissak,r.i.,and Drasgow,F.(1982). Recovery of two and three parameter logistic item characterisic curves : Amonte Carlo Study. Applied Psychological Measurement, No.(6),PP(249-260).

Hwang ,Dae-Yeop.(2002). Classical test theory and item response theory :Analytical and empirical comparisons . Paper presented at the annual meeting of Southwest Educational ResearchAssosiation (Austin,tx, February 14-16,2002).

Kelley,T.L.(1939). Selectionof upper and lower groupsfor the validation of test items . Journal of Educational Psychology, No. 30,PP(17-24).

Linn,R.L.(1990). Has item response theory increased the validity of achievement test scores? .Applied Measurement in Education.Vol. (3) , No.(2), PP (115-141).

Lord, F.M.(1980).Applications of item response theory to practical testing problems. Hillsdale MJ:Lawrence Erlbaum Assosiates.

Lord, F.M. & Novick, M.R. (1968). Statistical theories of mental test scores. Reading, Massachusetts: Addison-Wesley.

MacDonald, P. & Paunonen, S.V. (2002). A Monte Carlo comparison of item and person statistics based on item response theory versus classical test theory. Educational and Psychological Measurement, Vol.(62),No.(6),PP(921-943).

Matlock,S.(1997). Basic concepts in item and test analysis . Retrieved July 17.2002.from Texas A & M University Web site: http://www.ericae.net/ft/tamu/Espy.htm

Stage, C. (1998). A comparison between item analysis based on item response theory and classical test theory: a study of the SweSAT Subtest READ. Educational Measurement No 30. Umeå, Sweden: University of Umeå, Department of Educational Measurement. From www.umu.se/edmeas/publikationer/index_eng.html .

Truskosky, Denise. Marie. (1999). An empirical examination of classical test theory and item response theory parameters: Implications for research and practice in small –and large- sample assessments. (Doctoral dissertation. Southern Illonois University at Carbondale). Dessertation Abstract International, DAI-B60/08, p.4297

Wright,B.D.(1967).sample-free test calibration and person measurement.. Retrived March 16 , 2002 . from htt://www.rasch.org/ rasch transactions .

Wright B.D. (1977) Solving measurement problems with the Rasch model. Journal of Educational Measurement,Vol.(14),No.(2),PP(97-116).

Wright,B.& Lincare,John M.(1998). Auser Guide to Bigsteps/Winsteps: Rasch- Model Computer Program..Chicago,IL:MSEA Press

Wright,D.&Stone,M.(1979).Best test design: Ahand book for Rasch Measurement. Chicago:MESA Press

الملحق رقم (1)

الاختبار الوطني للمرحلة الأساسية في مبحث العلوم للصف التاسع

بسم الله الرحمن الرحيم

وزارة التربية والتعليم

المديرية العامة للامتحانات والاختبارات

مديرية الاختبارات

الاختبار الوطني للمرحلة الأساسية

للعام الدراسي ٢٠٠٠/٢٠٠١

المبحث : العلوم

لاستعمال الطالب/الطالبة

الاسم : ....

الجنس : ذكر ☐ أنثى ☐

تاريخ الولادة : / / ١٩

المدرسة : ....

المديرية : ....

لاستعمال الحاسوب

رمز الطالب :	☐☐☐☐
رمز المادة :	٦
رمز الجنس :	☐
رمز المديرية :	☐☐
رمز المدرسة :	☐☐☐☐

تعليمات الاختبار :-

١- اكتب المعلومات في المستطيل المخصص لاستعمال الطالب/الطالبة وعلى نموذج الإجابة.

٢- عدد أسئلة الاختبار (٩) وعدد صفحاته (١٠).

٣- يتكون هذا الاختبار من نوعين من الأسئلة :

النوع الأول : أسئلة الاختيار من متعدد وعددها (٢٠) فقرة لكل فقرة أربعة بدائل، واحد منها فقط صحيح. ضع رمز البديل الصحيح على الجزء المخصص لذلك في الصفحة الأولى.

النوع الثاني : أسئلة مقالية وعددها (٨). أجب عن كل منها في الفراغ المخصص بعد نهاية كل سؤال.

٤- مدة الاختبار ساعة ونصف.

نموذج الإجابة

اسم الطالب :	المدرسة :	المديرية :	المبحث : العلوم
رمز الطالب :	رمز المدرسة :	رمز المديرية :	رمز المبحث : ٦

• الجزء المخصص لإجابة الطالب على فقرات الاختيار من متعدد : العلامة الكلية : (100)

السؤال الأول : ( ) علامة

رقم الفقرة	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	المجموع
رمز الإجابة																										
العلامة																										

المصحح الأول : المصحح الثاني :

• الجزء المخصص لرصد علامات اسئلة الاختبار المقاليـــــــة :

رقم السؤال	رمز الفرع	العلامة المستحقة	رمز الفرع	العلامة المستحقة	المصححون
2	أ		ج		
	ب		د		
3	أ		ج		
	ب		د		
4	أ		ج		
	ب		د		
5	أ		ج		
	ب		د		
6	أ		ج		
	ب		د		
7	أ		ج		
	ب		د		
8	أ		ج		
	ب		د		
9	أ		ج		
	ب		د		
المجموع					

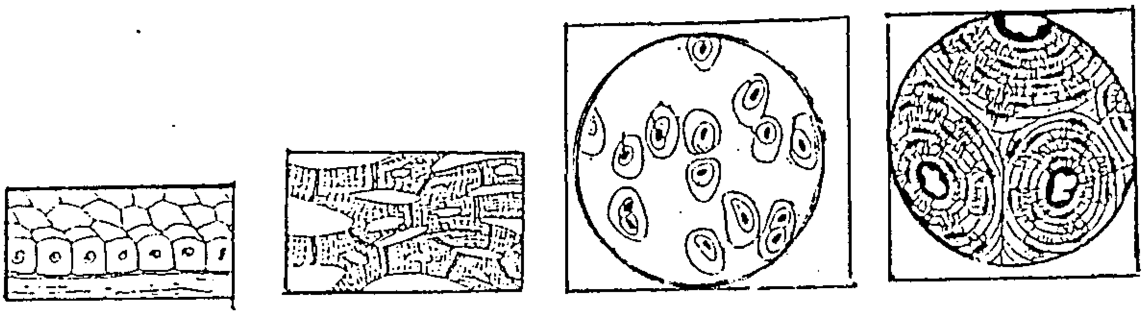
اسم مدخل البيانات : اسم مراجعها :

السؤال الأول : يتكون هذا السؤال من (٢٠) فقرة من نوع الاختيار من متعدد. ضع رمز الإجابة الصحيحة على النموذج المرفق في المكان المخصص لذلك. (٢٠ علامة)

١- أي الأعضاء الآتية لا يوجد في التجويف البطني :

أ. الأمعاء ب. البنكرياس جـ. الكلية د. الرئتان

٢- أحد الأنسجة الآتية يمثل نسيجاً عضلياً قلبياً :

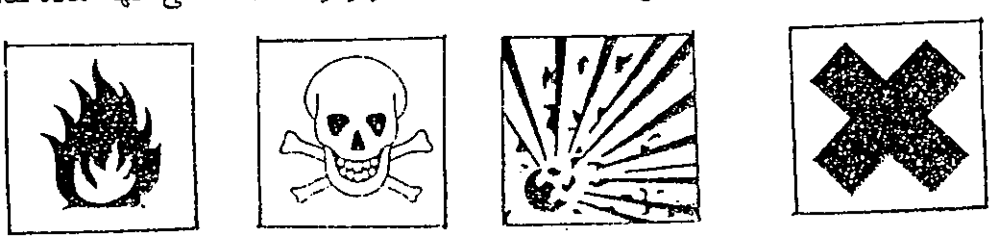


أ. ب. جـ. د.

٣- المادة الغذائية التي يصعب هضمها عند استئصال الحوصلة الصفراوية (المرارة) هي :

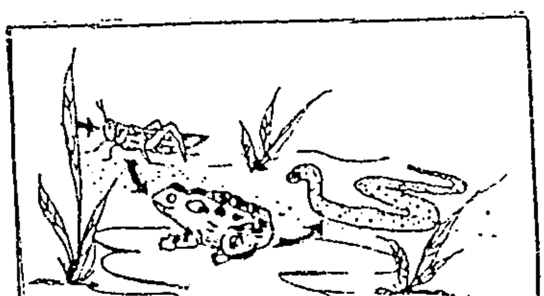
أ. السكريات ب. البروتين جـ. الدهون د. السيليلوز

٤- أي الإشارات الآتية توضع على عبوة المادة الكيميائية للدلالة على أنها مادة سامة :



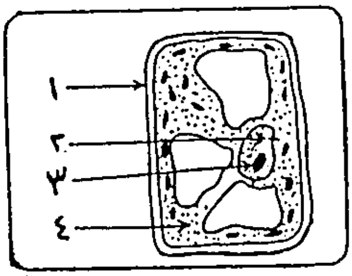
أ. ب. جـ. د.

٥- في الشكل أدناه والذي يمثل سلسلة غذائية، ما الحيوان يعد مستهلكاً أولاً :



أ. الضفدع ب. الجراد جـ. الأفعى د. النبات

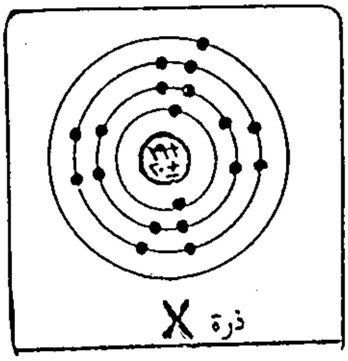
٦- رقم التركيب في الشكل المجاور الذي يدل على أن الخلية هي خلية نباتية :



أ. (١) ب. (٢)

جـ. (٣) د. (٤)

٧- يمثل الرسم المجاور بناء لذرة العنصر X ، إحدى العبارات الآتية تعتبر صحيحة بالنسبة للذرة X :



أ. العدد الكتلي = ٢٠ جـ. عدد النيوترونات = ١٩

ب. العدد الذري = ٢٠ د. عدد البروتونات = ١٩

٨- عند تفاعل العنصر $_{13}X^{27}$ مع العنصر $_{8}Y^{16}$ ،فإن الصيغة الكيميائية للمركب الناتج:

أ. XY_2 ب. X_2Y جـ. X_2Y_3 د. X_3Y_2

٩- أحد التغيرات الآتية يمثـل تفاعلاً كيميائياً :

أ. ذوبان ملح الطعام في الماء جـ. تحلل الماء

ب. صهر النحاس د. إضافة الماء إلى حمض H_2SO_4

١٠- أحد المركبات الآتية يمثل أكسيداً لا فلزياً :

أ. CaO ب. Na_2O جـ. SO_2 د. MgO

١١- حضر أحمد محلولاً ملحياً بإذابة ٥ غ من ملح في ٥٠ مل من الماء، وأراد أن يحضـر محلولاً ملحياً تركيزه نصف تركيز المحلول السابق فعليه إذابة :

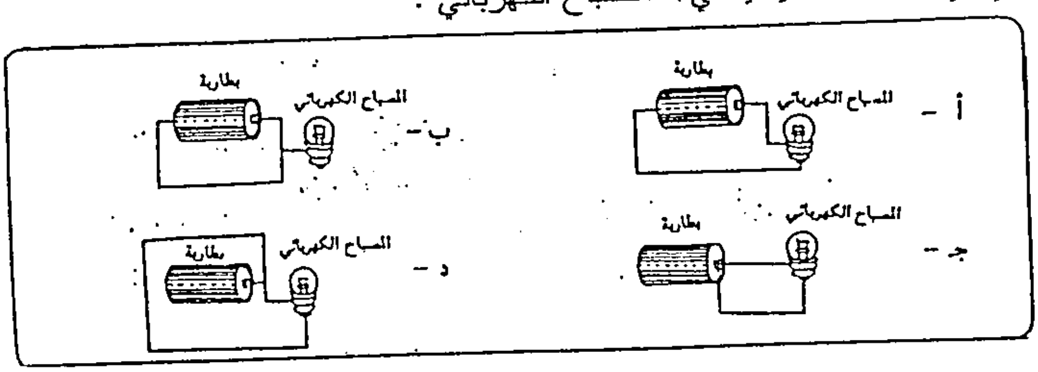
أ. ١٠ غ من الملح في ١٠٠مل من الماء ب. ١٠غ من الملح في ٥٠مل من الماء

جـ. ٥ غ من الملح في ١٠٠مل من الماء د. ٥غ من الملح في ٢٥مل من الماء

١٢- أحد الأجهزة الآتية من أدوات قياس سرعة الرياح :

أ. الباروميتر ب. الأنيموميتر جـ. الفولتميتر د. الأميتر

١٣- في أي الحالات الآتية يضيء المصباح الكهربائي :

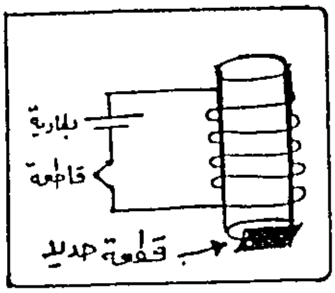


١٤- وضع جسم أمام عدسة محدبة الوجهين فتكون له خيال (مكبر، معتدل، وهمي)، فـــإن الجسم في هذه الحالة يقع على بعد:

أ. س > ٢ع　　　　ب. س = ٢ ع

جـ. ٢ ع > س > ع　　　　د. س < ع

ملحوظة : (س) بعد الجسم عن مركز العدسة، (ع) البعد البؤري للعدسة.



١٥- الشكل المجاور يمثل دارة ملف حلزوني مغلقة.
إذا تم فتح دارة هذا الملف فإن القطعة الحديدية:

أ. تسقط إلى أسفل　　　　ب. تبقى مجذوبة للملف

جـ. تدور　　　　د. تهتز

١٦- إذا كانت كتلة جسم تساوي (٤٠ كغ) وتسارع الجاذبية الأرضية يساوي ١٠ م/ث^{٢} فإن وزن الجسم (بالنيوتن) يساوي :

أ. ٤　　　　ب. ٣٠　　　　جـ. ٤٠٠　　　　د. ٤,٠

١٧- تعرضت قطعتا حديد وألومنيوم معا (متساويتان في الكتلة) إلى حرارة الشمس لمدة ١٠ دقائق، فارتفعت درجة حرارتهما فوجد أن القطعتين متساويتان في :

أ. درجة الحرارة　　　　جـ. كمية الحرارة المكتسبة

ب. الحرارة النوعية　　　　د. السعة الحرارية

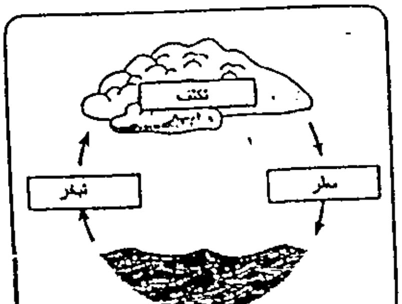
١٨- الغرانيت نوع من أنواع الصخور :

أ. الرسوبية　　　　ب. المتحولة　　　　جـ. النارية　　　　د. البازلتية

١٩- يبين الرسم المجاور دورة الماء في الطبيعة،
ما مصدر الطاقة بدورة الماء ؟

أ. القمر　　　　ب. الشمس

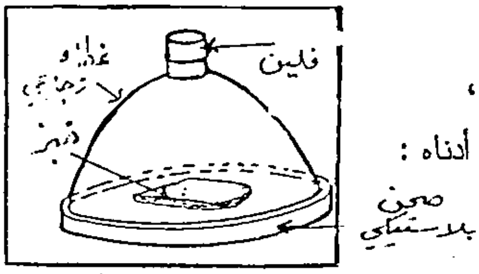
جـ. المد والجزر　　　　د. الرياح



٢٠- الطبقة في الغلاف الجوي التي تحدث فيها التقلبات الجوية، وينتج فيها التكاثف ومنها يتم الهطول هي الغلاف :

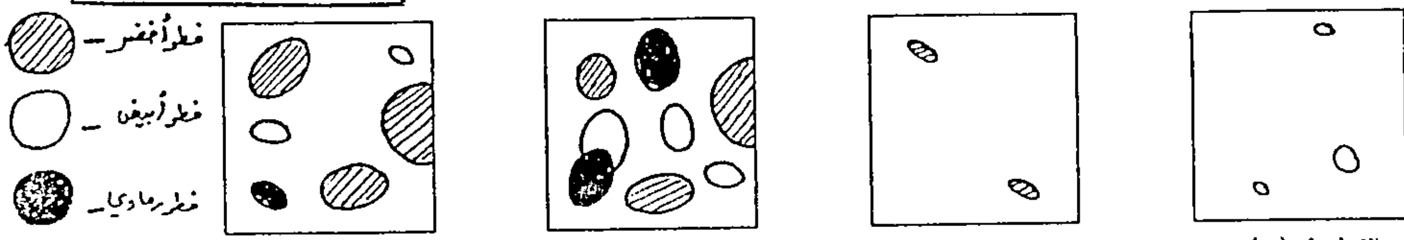
أ. الخارجي　　　　ب. المتوسط　　　　جـ. المناخي　　　　د. الطبقي

القطعة	درجة الحرارة س	الظروف
١	٣٠	جاف
٢	٢٠	جاف
٣	٣٠	رطب
٤	٢٠	رطب



<u>السؤال الثاني :</u> (١١ علامة)

أ- تنمو الفطريات على الخبز فتؤدي إلى تعفنه ويصبح غير صالح للأكل، أراد أحد الطلاب دراسة كيفية خزن الخبز دون تعفنه. فقام بتجفيف (٤) قطع من الخبز من نفس الرغيف ، وقد عولجت القطع كما هو مبين في الجدول المجاور .وخلال التجربة حفظت كل قطعة أسفل غطاء بلاستيكي كما في الشكل المجاور، وبعد ١٠ أيام ظهر نمو الفطريات على كل قطعة خبز كما في الشكل أدناه :



القطعة (٤) القطعة (٣) القطعة (٢) القطعة (١)

من خلال ما سبق، أجب عما يأتي :

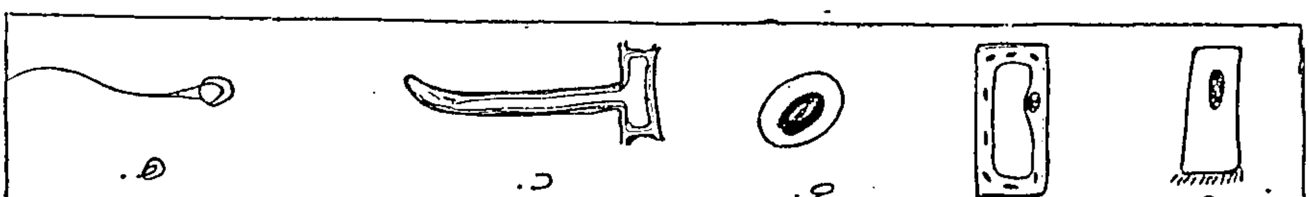
١. لماذا تم تجفيف الخبز في بداية التجربة؟..................

٢. لماذا وضع غطاء بلاستيكي شفاف على العينات الأربع السابقة الذكر؟.........

٣. أي الفطريات ينمو بشكل أفضل على درجات الحرارة المنخفضة؟...............

ب- مما سبق أي الظروف أفضل لحفظ الخبز ؟............................

ب- يشير الشكل الآتي إلى عدد من أنواع من الخلايا ، وكل منها يقوم بوظيفة مختلفة : (٤ علامات)



١. ما اسم الخلية (جـ) . ...

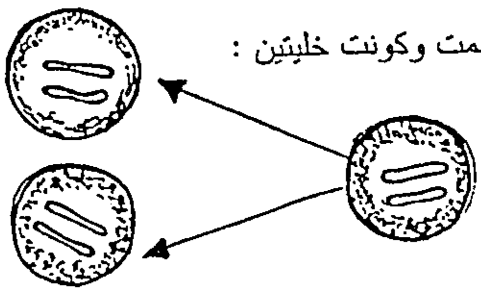
٢. ما رمز الخلية في الشكل أعلاه التي تقوم بما يأتي :

- البناء الضوئي. ....

- طرد الأجسام الغريبة إلى خارج الجسم. ....

٣. اكتب اسم الجزء من النبات الذي توجد به الخلية (د). ....

جـ- في الشكل الآتي خلية انقسمت وكونت خليتين :   (٣ علامات)



١. ما نوع الإنقسام الذي حدث في الخلية وأدى إلى تكوين الخليتين؟ ....

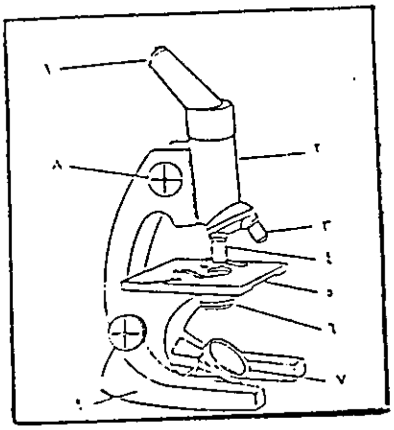
٢. كم عدد الكروموسومات في كل خلية من الخلايا الناتجة؟ ....

٣. ما أهمية هذا النوع من الإنقسام؟ ....

....

السؤال الثالث :   (١١ علامة)

أ- يمثل الشكل المجاور مجهراً ضوئياً مركباً، أجب عما يأتي :



٢. ما وظيفة الجزء رقم (٨) ؟ ....

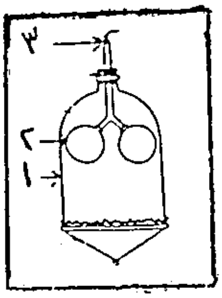
٢. اكتب أسماء الأجزاء المشار إليها بالأرقام (١،٦، ٧).

(١) ....، (٦) ....، (٧) ....

٣. ما قوة تكبير المجهر إذا كان قوة تكبير العدسة العينية (١٠ مرات) والشيئية (١٠٠ مرة) ؟ ....

(٥ علامات)

ب- يمثل الشكل المجاور نموذجاً مماثلاً لتركيب الجهاز التنفسي :



(٣ علامات)

١. ماذا تمثل الأرقام (١،٢) في الشكل ؟

(١) ....، (٢) ....

٢. هل يمثل الشكل عملية شهيق أم زفير؟ ....

جـ- لديك (٣) أنابيب اختبار يحتوي كل منها على سائل شفاف، أحد هذه الأنابيب يحتوي على مادة بروتينية، كيف تستطيع تحديد الأنبوب الذي يحتوي عليها. (٣ علامات)

....

....

....

السؤال الرابع : (١١ علامة)

أ- قام بسام بغمس أوراق ٣ أنواع من الورد : (صفراء، حمراء، بنفسجية) لمعرفة تغيرلون كل منها في الماء المقطر، أو المحلول الحمضي، أو المحلول القاعدي، ودوّن ملاحظاته في الجدول المجاور، اعتماداً على هذه المعلومات، اختر نوعاً واحداً من أوراق الورد يكون مناسباً أكثر ككاشف عن المحاليل : المتعادلة والحمضية والقاعدية معاً. ....

فسر إجابتك: ....

لون ورقة الورد	في محلول حمضي	في الماء المقطر	في محلول قاعدي
اصفر	اصفر	اصفر	اصفر
بنفسجي	احمر	بنفسجي	أزرق
احمر	احمر	احمر	اخضر

ب- مستعيناً بالجدول المجاور، أجب عما يأتي مستخدماً الرموز : (أ، ب، جـ، د، هـ، و).

١. هناك ذرتان تمثلان نفس العنصر وهما : .....، .....

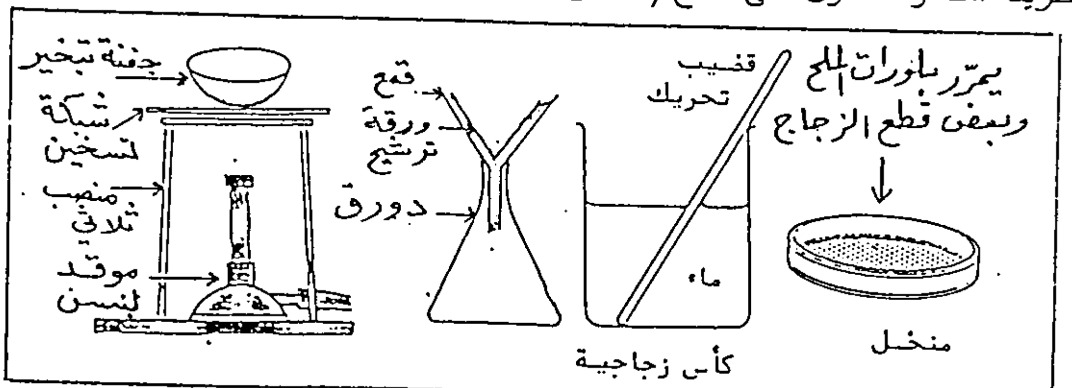
٢. أي الرموز يمثل أيوناً موجباً؟ ..... .

٣. أي ذرة (أو أيون) العدد الكتلي لها يساوي ٢٣ ؟ ..... .

٤. أي منها يمثل عنصراً نبيلاً (خاملاً) ؟ ..... .

الذرة أو الأيون	عدد البروتونات	عدد النيوترونات	عدد الإلكترونات
أ	١٠	١٠	١٠
ب	١١	١٢	١١
جـ	١٣	١٤	١٠
د	١٧	٢٠	١٧
هـ	٩	١٠	١٠
و	١٧	١٨	١٧

جـ. تناول رامي زجاجة تحتوي ملح نترات البوتاسيوم (KNO_3) من خزانة المواد في المختبر فسقطت على الأرض وانكسرت، اقترح طريقة لفصل قطع الزجاج المتناثر عن الملح بطريقة آمنة والحصول على الملح.(استعن بما تراه مناسباً من الأدوات المبينة في الرسم).



- -

- -

- -

السؤال الخامس : (١٠ علامات)

أ- وصف أحمد طريقة تحضيره لبلورات من كبريتات النحاس (II) فقال :

الخطوة الأولى : أضفت كمية من أكسيد النحاس CuO إلى كأس يحتوي محلول حمض الكبريتيك الساخن وحركت المزيج جيداً حتى أصبح المحلول أزرق صافياً، ثم أضفت كمية أخرى من أكسيد النحاس إلى المحلول حتى وصل إلى درجة فوق الإشباع.

الخطوة الثانية : قمت بترشيح المحلول ووضعته في كأس زجاجية وظهر راسب أسود على ورقة الترشيح.

الخطوة الثالثة : تركت المحلول الراشح ليبرد تدريجياً فلاحظت تكون بلورات زرقاء في قاع الكأس الزجاجية.

من المعلومات السابقة أجب عن الأسئلة الآتية :

١- ما لون كبريتات النحاس (II) ؟ ....

٢- اكتب معادلة التفاعل الذي حدث في الكأس الزجاجية (الخطوة الأولى) بالكلمات.

....

٣- لماذا قام أحمد بترشيح المحلول (الخطوة الثانية)؟

....

ب- يبين الجدول الآتي تفاعل بعض الفلزات مع الماء البارد ومحلول حمض HCl المخفف :

العنصر	الخارصين	البلاتين	البوتاسيوم	النيكل
التفاعل مع الماء البارد	لا يتفاعل	لا يتفاعل	ينصهر ويظهر لهب ويحدث انفجار أحياناً	لا يتفاعل
التفاعل مع محلول HCl المخفف	يتفاعل ويذوب ببطء وتظهر فقاعات غازية	لا يتفاعل	لا يمكن إجراء التجربة بشكل آمن في المختبر	تظهر فقاعات من الغاز إذا تم تسخين محلول HCl

اعتماداً على المعلومات الواردة في الجدول، أجب عما يأتي :

١. رتب العناصر الواردة في الجدول حسب نشاطها، من الأكثر نشاطاً إلى الأقل نشاطاً.

....

٢. ما اسم الغاز الذي يتكون عند تفاعل الخارصين مع محلول حمض HCl المخفف؟......

٣. لا ينبغي إجراء تجربة تفاعل البوتاسيوم مع محلول حمض HCl المخفف في المختبر، بيّن لماذا يكون إجراء هذه التجربة خطيراً؟

....

....

السؤال السادس :   (٩ علامات)

أ- فحصت ٤ عينات من الماء من مناطق مختلفة في الأردن وذلك بإضافة تدريجية لمحلول صابوني إلى ٢٥مل من كل عينة مع التحريك وذلك حتى تتكون رغوة دائمة وكانت النتائج كما هو مبين في الجدول المجاور، مستعيناً بالنتائج أجب عما يأتي :

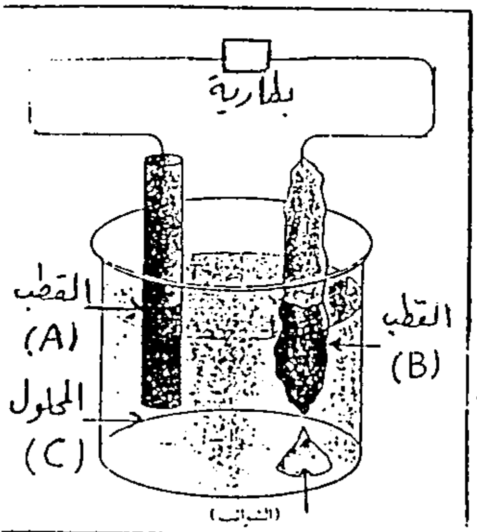
عينة الماء الحجم (٢٥مل)	حجم محلول الصابون المضاف (مل) حتى تتكون الرغوة: قبل الغلي	بعد الغلي
أ	١٢	١
ب	١٦	٨
جـ	١٢	١٢
د	١٤	١
جـ بعد التقطير	....	....

١. حدد العينة (أو العينات) التي تمثل العسر المؤقت. ......

٢. حدد العينة (أو العينات) التي تمثل مزيجاً من العسر الدائم والعسر المؤقت......

٣. إملأ الفراغين الموجودين في الجدول.

ب- يمثل الرسم المجاور طريقة تنقية النحاس بالتحليل الكهربائي، اعتماداً على الرسم، أجب عما يأتي :



١. ما اسم المادة المكونة للقطب (٨)؟...........

٢. ما اسم المحلول (C)؟ ....

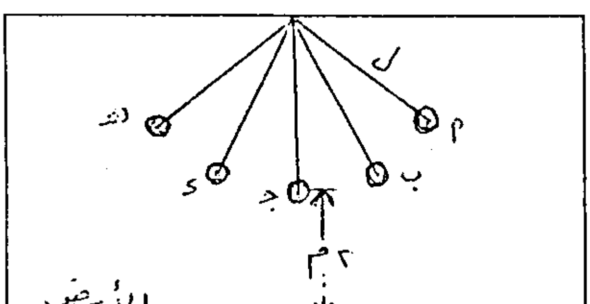
٣. في أي قطبي البطاربة يتم وصل القطب (B)؟ ....

٤. اكتب المعادلة الأيونية التي تحدث على القطب (٨).

....

السؤال السابع :   (٩ علامات)

أ- الشكل المجاور يمثل بندولاً طوله ل، أفلت من النقطة (أ) ليصل إلى النقطة (هـ) بالاعتماد على الشكل والمعلومات المثبتة عليه، أجب عما يأتي :

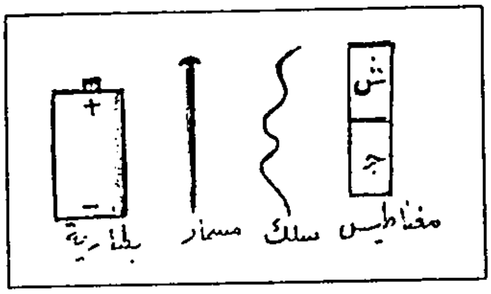


١. عند أي المواقع يمتلك البندول أكبر طاقة وضع ......،أكبر طاقة حركة ....

٢. عند أي المواقع يمتلك البندول طاقة وضع وطاقة حركة في آن واحد ....

٣. احسب طاقة وضع البندول عند النقطة (جـ) بالنسبة للأرض، إذا علمت أن كتلة البندول ٢ كغ وتسارع الجاذبية ١٠ م/ث^{٢}.

....



ب- وضح كيف تصنع من المسمار مغناطيساً
باستخدام البطارية تارة ودون استخدام البطارية
تارة أخرى مستخدماً الأدوات الواردة في الشكل المجاور.

....

....

<u>السؤال الثامن :</u> (٩ علامات)

أ- علل لكل مما يأتي تعليلاً وافياً : (٤ علامات)

١. لا تملأ زجاجات المشروبات الغازية بشكل كامل.

....

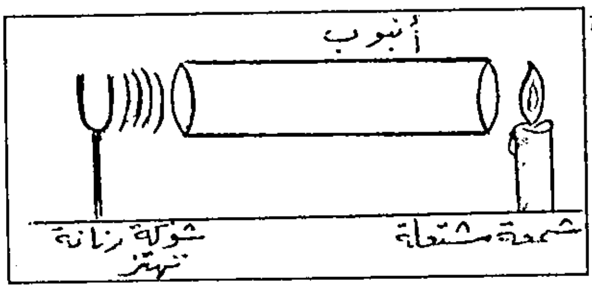
....

٢. تبقى درجة حرارة الثلج والماء المنصهر داخل وعاء خلال عملية الإنصهار ثابتة.

....

....

ب- أنبوب مفتوح الطرفين وضعت شمعة مشتعلة (٥ علامات)
مقابل أحد طرفيه وفي الطرف الآخر شوكة رنانة مهتزة
فسر ماذا يحدث للهب الشمعة في الحالتين الآتيتين:



١. عند وجود الهواء داخل الأنبوب.

....

٢. عند تفريغ الأنبوب من الهواء.

....

السؤال التاسع : (١٠ علامات)

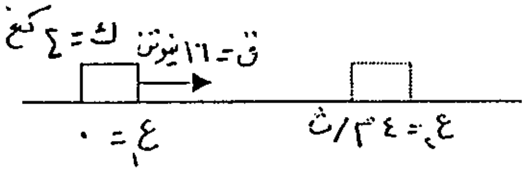
أ- سحب جسم كتلته ٤كغ من السكون بقوة أفقية قدرها ١٦نيوتن، فإذا أصبحت سرعته ٤م/ث بعد مضي ن ثانية، اعتبر تسارع الجاذبية ١٠م/ث٢ .

احسب : ١. تسارع الجسم.

....

٢. مدة تأثير القوة.

....



ب- القائمة (أ) تمثل مجموعة من الكميات والمجموعة (ب) تمثل الأجهزة أو الأدوات التـــي تقيس هذه الكميات. صل بين الكمية من العمود (أ) وما يناسبها من العمود (ب).

العمود (أ)	العمود (ب)
فرق الجهد	السيزموجراف
شدة التيار	الأنيموميتر
الكتلة	ميزان ذو الكفين
شدة الزلزال	الأميتر
	الفولتميتر

انتهت الأسئلة

المقدمة

يعد قياس مستوى الطلبة على اختبار ضبط نوعية التعليم للمرحلة الأساسية معيارا يمكن من خلاله الحكم على فاعلية العملية التربوية بعناصرها المختلفة ، وتحديد نقاط القوة وتعزيزها ، ونقاط الضعف لتشخيصها ومعالجتها . ولتحقيق ذلك يأتي مشروع هذا الاختبار للحكم على مدى امتلاك الطلبة للمهارات الأساسية بعد أن أمضوا عددا من السنوات على مقاعد الدراسة .

تعريف الاختبار الوطني

اختبار تعده وزارة التربية والتعليم (المديرية العامة للامتحانات) بهدف قياس مدى امتلاك طلبة الفئة المستهدفة للمهارات الأساسية .

خطوات بناء الاختبار

أولا : بناء الاختبار

١. تحديد فرق بناء الأسئلة للمباحث الستة الآتية :
التربية الإسلامية ، واللغة العربية ، واللغة الإنجليزية ، والرياضيات ، والعلوم ، والاجتماعيات للصف التاسع .

٢. مهام فرق بناء الأسئلة

أ. تحليل محتوى الكتب الدراسية للصفوف من الأول حتى التاسع .

ب. إعداد جدول مواصفات الاختبار .

ج. بناء اختبارين متكافئين .

د. طباعة ونسخ الاختبارين .

ثانيا : تجريب الاختبار

١. تجريب الاختبارين على عينة من طلبة الصف التاسع في مناطق مختلفة ولكلا الجنسين .

٢. تصحيح الاختبارين وتفريغ نتائجهما .

٣. التحليل الإحصائي لنتائج الاختبارين لإيجاد فاعلية البدائل ، ومعامل صعوبة الفقرات وتمييزها ، ومتوسط الأداء على كل فقرة .

٤. اختيار الفقرات الملائمة اعتمادا على نتائج التحليل الإحصائي ، وإعداد اختبار نهائي منسجم وجدول المواصفات .

ثالثا : اختيار عينة الدراسة

١. اختيار عينة عشوائية منظمة تشمل (٥%) من طلبة الصف التاسع الأساسي في كل مديريات التربية والتعليم مع مراعاة المتغيرات الآتية :
الجنس ، وعدد طلبة المديرية ، والسلطة التعليمية المشرفة ، المنطقة السكنية (ريف ، مدينة) .

٢. اختيار المدارس التي سيتم فيها تطبيق الاختبار بطريقة عشوائية .

٣. اختيار طلبة الشعبة الواحدة جميعهم أما إذا قل عدد طلبة الشعبة عن العدد المطلوب فيتم إكمال العدد من شعبة أخرى وبحسب الحروف الهجائية .

أهداف الاختبار

إن الاختبار لا يهدف إلى تقييم الطلبة وإعطائهم علامات للترفيع أو النجاح أو الانتقال من مرحلة دراسية إلى أخرى وإنما يهدف إلى :

- تزويد المعنيين بالتربية والتعليم بتحصيل الطلبة فيما يتعلق بالمهارات الأساسية التي يجب أن يتقنها الطلبة .
- تزويد المعنيين بالتربية والتعليم بالمعلومات الضرورية المتعلقة بتحصيل الطلبة في المدارس ومديريات التربية والتعليم ، وبذلك يعتبر تقويما ختاميا.
- تحديد جوانب القوة في تعلم الطلبة لمختلف المهارات وتحديد الخطوات الملائمة لتعلمهم المستقبلي ، وبذلك يعتبر تقويما تكوينا .
- تشخيص جوانب الضعف لدى الطلبة وتحديد مصادرها ووضع الخطط العلاجية المناسبة ، وبذلك يعتبر تقويما تشخيصيا.
- من جوانب العملية تقييم البرنامج التعليمي ككل : وذلك بتحديد مواطن الخلل فيما إذا كانت مرتبطة بالمنهاج أو طريقة التدريس أو أي جانب آخر من جوانب العملية التربوية .

الملحق رقم (2)

أهداف الاختبار وخطوات بنائه واجراءآت تطبيقه

الملحق رقم (3)

قدرات الطلبة المفحوصين على الاختبار وفق نموذج راش باستخدام برنامج Winsteps

INPUT: 5520 PERSONS, 20 ITEMS ANALYZED: 5520 PERSONS, 20 ITEMS

SCORE العلامة	MEASURE القدرة	S.E. خطا معياري	SCORE العلامة	MEASURE القدرة	S.E. خطأ معياري	SCORE العلامة	MEASURE القدرة
0	-4.28E	1.84	7	-0.65	0.48	14	0.88
1	-3.03	1.03	8	-0.43	0.47	15	1.15
2	-2.28	0.75	9	-0.21	0.46	16	1.45
3	-1.80	0.63	10	0.00	0.46	17	1.81
4	-1.45	0.57	11	0.21	0.46	18	2.28
5	-1.15	0.53	12	0.42	0.47	19	3.05
6	-0.89	0.50	13	0.65	0.48	20	4.30E

ملحق (4)

اختبار (t-test) لأداء الطلبة على الاختبارين السهل والصعب وفق النظرية الكلاسيكية

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
EASY	5520	6.7165	1.6617	2.237E-02
HARD	5520	5.3170	1.6293	2.193E-02

T-Test

One-Sample Test

	Test Value = 0					
	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference Lower	95% Confidence Interval of the Difference Upper
EASY	300.303	5519	.000	6.7165	6.6726	6.7603
HARD	242.456	5519	.000	5.3170	5.2740	5.3600

ملحق رقم 5

قدرات الطلبة على الاختبار السهل

INPUT: 5520 PERSONS, 10 ITEMS ANALYZED: 5329 PERSONS, 10 ITEMS .

PERSON STATISTICS: MEASURE ORDER

رقم الطالب ENTRY NUMBER	العلامة الخام RAW SCORE	COUNT	القدرة MEASURE	ERROR	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD	PTBIS CORR.	نمط الاستجابة PERSON
1	10	10	3.52	1.85	MAXIMUM ESTIMATED MEASURE					1111111111
2	10	10	3.52	1.85	MAXIMUM ESTIMATED MEASURE					1111111111
3	10	10	3.52	1.85	MAXIMUM ESTIMATED MEASURE					1111111111
4	10	10	3.52	1.85	MAXIMUM ESTIMATED MEASURE					1111111111
5	10	10	3.52	1.85	MAXIMUM ESTIMATED MEASURE					1111111111
6	10	10	3.52	1.85	MAXIMUM ESTIMATED MEASURE					1111111111
7	10	10	3.52	1.85	MAXIMUM ESTIMATED MEASURE					1111111111
8	10	10	3.52	1.85	MAXIMUM ESTIMATED MEASURE					1111111111
9	10	10	3.52	1.85	MAXIMUM ESTIMATED MEASURE					1111111111
10	10	10	3.52	1.85	MAXIMUM ESTIMATED MEASURE					1111111111
11	10	10	3.52	1.85	MAXIMUM ESTIMATED MEASURE					1111111111
12	10	10	3.52	1.85	MAXIMUM ESTIMATED MEASURE					1111111111
13	10	10	3.52	1.85	MAXIMUM ESTIMATED MEASURE					1111111111
14	10	10	3.52	1.85	MAXIMUM ESTIMATED MEASURE					1111111111
15	10	10	3.52	1.85	MAXIMUM ESTIMATED MEASURE					1111111111
16	10	10	3.52	1.85	MAXIMUM ESTIMATED MEASURE					1111111111
17	10	10	3.52	1.85	MAXIMUM ESTIMATED MEASURE					1111111111
18	10	10	3.52	1.85	MAXIMUM ESTIMATED MEASURE					1111111111
19	10	10	3.52	1.85	MAXIMUM ESTIMATED MEASURE					1111111111
20	10	10	3.52	1.85	MAXIMUM ESTIMATED MEASURE					1111111111
21	10	10	3.52	1.85	MAXIMUM ESTIMATED MEASURE					1111111111
22	10	10	3.52	1.85	MAXIMUM ESTIMATED MEASURE					1111111111
23	10	10	3.52	1.85	MAXIMUM ESTIMATED MEASURE					1111111111
24	10	10	3.52	1.85	MAXIMUM ESTIMATED MEASURE					1111111111
25	10	10	3.52	1.85	MAXIMUM ESTIMATED MEASURE					1111111111
26	10	10	3.52	1.85	MAXIMUM ESTIMATED MEASURE					1111111111
27	10	10	3.52	1.85	MAXIMUM ESTIMATED MEASURE					1111111111
28	10	10	3.52	1.85	MAXIMUM ESTIMATED MEASURE					1111111111
29	10	10	3.52	1.85	MAXIMUM ESTIMATED MEASURE					1111111111
30	10	10	3.52	1.85	MAXIMUM ESTIMATED MEASURE					1111111111
31	10	10	3.52	1.85	MAXIMUM ESTIMATED MEASURE					1111111111
32	10	10	3.52	1.85	MAXIMUM ESTIMATED MEASURE					1111111111
33	10	10	3.52	1.85	MAXIMUM ESTIMATED MEASURE					1111111111
34	10	10	3.52	1.85	MAXIMUM ESTIMATED MEASURE					1111111111
35	10	10	3.52	1.85	MAXIMUM ESTIMATED MEASURE					1111111111
36	10	10	3.52	1.85	MAXIMUM ESTIMATED MEASURE					1111111111
37	10	10	3.52	1.85	MAXIMUM ESTIMATED MEASURE					1111111111
38	10	10	3.52	1.85	MAXIMUM ESTIMATED MEASURE					1111111111
39	10	10	3.52	1.85	MAXIMUM ESTIMATED MEASURE					1111111111
40	10	10	3.52	1.85	MAXIMUM ESTIMATED MEASURE					1111111111
41	10	10	3.52	1.85	MAXIMUM ESTIMATED MEASURE					1111111111
42	10	10	3.52	1.85	MAXIMUM ESTIMATED MEASURE					1111111111
43	10	10	3.52	1.85	MAXIMUM ESTIMATED MEASURE					1111111111
44	10	10	3.52	1.85	MAXIMUM ESTIMATED MEASURE					1111111111
45	10	10	3.52	1.85	MAXIMUM ESTIMATED MEASURE					1111111111
46	10	10	3.52	1.85	MAXIMUM ESTIMATED MEASURE					1111111111
47	10	10	3.52	1.85	MAXIMUM ESTIMATED MEASURE					1111111111
48	10	10	3.52	1.85	MAXIMUM ESTIMATED MEASURE					1111111111
49	10	10	3.52	1.85	MAXIMUM ESTIMATED MEASURE					1111111111
5472	3	10	-0.87	0.70	0.92	-0.3	0.89	-0.4	0.41	0000011001

5473	3	10	-0.87	0.70	1.12	0.4	1.15	0.5	-0.20	0001101000
5474	3	10	-0.87	0.70	1.03	0.1	1.07	0.2	0.01	0011100000
5475	3	10	-0.87	0.70	1.04	0.1	1.07	0.2	0.00	1010100000
5476	3	10	-0.87	0.70	0.76	-0.9	0.70	-1.1	0.90	0010011000
5477	3	10	-0.87	0.70	0.92	-0.3	0.89	-0.4	0.41	0000011001
5478	3	10	-0.87	0.70	0.94	-0.2	0.91	-0.3	0.36	0100011000
5479	3	10	-0.87	0.70	1.05	0.2	1.08	0.2	-0.01	1000010001
5480	3	10	-0.87	0.70	1.12	0.4	1.15	0.5	-0.22	1000101000
5481	3	10	-0.87	0.70	0.91	-0.3	0.89	-0.4	0.43	0010101000
5482	3	10	-0.87	0.70	0.91	-0.3	0.89	-0.4	0.43	0010101000
5483	3	10	-0.87	0.70	1.00	0.0	0.99	0.0	0.18	0000010101
5484	3	10	-0.87	0.70	0.83	-0.6	0.80	-0.7	0.66	0010010100
5485	3	10	-0.87	0.70	0.95	-0.2	0.95	-0.2	0.30	1010001000
5486	3	10	-0.87	0.70	0.76	-0.9	0.70	-1.1	0.90	0010011000
5487	3	10	-0.87	0.70	1.01	0.1	1.01	0.0	0.12	0000110001
5488	3	10	-0.87	0.70	1.10	0.3	1.12	0.4	-0.14	0001001100
5489	2	10	-1.42	0.80	1.15	0.3	1.32	0.6	-0.43	0001100000
5490	2	10	-1.42	0.80	0.86	-0.3	0.73	-0.6	0.64	0010001000
5491	2	10	-1.42	0.80	1.15	0.3	1.32	0.6	-0.43	0001100000
5492	2	10	-1.42	0.80	0.98	0.0	1.02	0.1	0.15	1010000000
5493	2	10	-1.42	0.80	1.11	0.2	1.20	0.4	-0.23	0001000010
5494	2	10	-1.42	0.80	1.16	0.3	1.33	0.6	-0.44	1000100000
5495	2	10	-1.42	0.80	1.15	0.3	1.31	0.6	-0.41	1000000001
5496	2	10	-1.42	0.80	0.97	-0.1	1.01	0.0	0.16	0011000000
5497	2	10	-1.42	0.80	1.07	0.1	1.14	0.3	-0.10	1000001000
5498	2	10	-1.42	0.80	0.86	-0.3	0.73	-0.6	0.64	0010001000
5499	2	10	-1.42	0.80	1.07	0.1	1.14	0.3	-0.10	1000001000
5500	2	10	-1.42	0.80	0.97	-0.1	1.01	0.0	0.16	0011000000
5502	2	10	-1.42	0.80	1.15	0.3	1.32	0.6	-0.43	0001100000
5503	2	10	-1.42	0.80	1.08	0.2	1.11	0.2	-0.10	0000100010
5504	2	10	-1.42	0.80	1.04	0.1	1.05	0.1	0.02	0100001000
5505	2	10	-1.42	0.80	1.04	0.1	1.05	0.1	0.02	0100001000
5506	2	10	-1.42	0.80	0.95	-0.1	0.92	-0.2	0.30	0010100000
5507	2	10	-1.42	0.80	1.00	0.0	1.05	0.1	0.10	1000010000
5508	2	10	-1.42	0.80	1.04	0.1	1.03	0.1	0.05	0000101000
5509	2	10	-1.42	0.80	0.86	-0.3	0.73	-0.6	0.64	0010001000
5510	2	10	-1.42	0.80	1.15	0.3	1.32	0.6	-0.43	0001100000
5511	2	10	-1.42	0.80	1.04	0.1	1.03	0.1	0.05	0000101000
5512	2	10	-1.42	0.80	1.16	0.3	1.33	0.6	-0.44	1000100000
5513	2	10	-1.42	0.80	1.11	0.2	1.18	0.3	-0.23	0000100100
5514	2	10	-1.42	0.80	0.99	0.0	0.92	-0.2	0.24	0000001010
5515	2	10	-1.42	0.80	0.97	-0.1	1.01	0.0	0.16	0011000000
5516	2	10	-1.42	0.80	0.95	-0.1	0.94	-0.1	0.27	0110000000
5517	1	10	-2.24	1.06	0.98	0.0	0.84	-0.2	0.26	0000001000
5518	1	10	-2.24	1.06	1.10	0.1	1.48	0.5	-0.39	1000000000
5519	1	10	-2.24	1.06	1.02	0.0	1.00	0.0	0.07	0000000010
5520	1	10	-2.24	1.06	1.09	0.1	1.45	0.4	-0.37	0001000000
MEAN	7.	10.	0.77	0.74	1.00	0.0	1.00	0.0		
S.D.	2.	0.	0.80	0.13	0.09	0.6	0.17	0.6		

ملحق رقم (6)

قيم قدرات الطلبة على الاختبار الصعب

INPUT: 5520 PERSONS, 10 ITEMS ANALYZED: 5511 PERSONS, 10 ITEMS.

PERSON STATISTICS: MEASURE ORDER

رقم

العلامةالخام الطالب القدرة

ENTRY NUMBER	RAW SCORE	COUNT	MEASURE	ERROR	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD	PTBIS CORR.	
1569	10	10	3.56	1.86	MAXIMUM ESTIMATED MEASURE					1111111111
1572	10	10	3.56	1.86	MAXIMUM ESTIMATED MEASURE					1111111111
4040	10	10	3.56	1.86	MAXIMUM ESTIMATED MEASURE					1111111111
4042	10	10	3.56	1.86	MAXIMUM ESTIMATED MEASURE					1111111111
4936	10	10	3.56	1.86	MAXIMUM ESTIMATED MEASURE					1111111111
4937	10	10	3.56	1.86	MAXIMUM ESTIMATED MEASURE					1111111111
1	9	10	2.26	1.07	1.11	0.1	1.35	0.4	-0.23	1111111101
128	9	10	2.26	1.07	0.74	-0.3	0.41	-0.8	0.93	0111111111
183	9	10	2.26	1.07	0.74	-0.3	0.41	-0.8	0.93	0111111111
187	9	10	2.26	1.07	1.11	0.1	1.40	0.4	-0.26	1111111110
659	9	10	2.26	1.07	1.07	0.1	1.13	0.1	-0.05	1110111111
661	9	10	2.26	1.07	0.74	-0.3	0.41	-0.8	0.93	0111111111
662	9	10	2.26	1.07	0.74	-0.3	0.41	-0.8	0.93	0111111111
663	9	10	2.26	1.07	1.08	0.1	1.15	0.2	-0.07	1111011111
664	9	10	2.26	1.07	1.11	0.1	1.34	0.3	-0.22	1111111011
666	9	10	2.26	1.07	1.08	0.1	1.18	0.2	-0.09	1111101111
667	9	10	2.26	1.07	1.11	0.1	1.35	0.4	-0.23	1111111101
668	9	10	2.26	1.07	1.02	0.0	0.88	-0.1	0.19	1011111111
1526	9	10	2.26	1.07	0.74	-0.3	0.41	-0.8	0.93	0111111111
1536	9	10	2.26	1.07	0.74	-0.3	0.41	-0.8	0.93	0111111111
1537	9	10	2.26	1.07	1.09	0.1	1.26	0.3	-0.16	1111110111
1542	9	10	2.26	1.07	1.07	0.1	1.13	0.1	-0.05	1110111111
1543	9	10	2.26	1.07	0.74	-0.3	0.41	-0.8	0.93	0111111111
1544	9	10	2.26	1.07	0.74	-0.3	0.41	-0.8	0.93	0111111111
1545	9	10	2.26	1.07	0.74	-0.3	0.41	-0.8	0.93	0111111111
1546	9	10	2.26	1.07	1.08	0.1	1.18	0.2	-0.09	1111101111
1548	9	10	2.26	1.07	1.07	0.1	1.13	0.1	-0.05	1110111111
1549	9	10	2.26	1.07	1.07	0.1	1.11	0.1	-0.04	1101111111
1555	9	10	2.26	1.07	1.02	0.0	0.88	-0.1	0.19	1011111111
1556	9	10	2.26	1.07	0.74	-0.3	0.41	-0.8	0.93	0111111111
1557	9	10	2.26	1.07	1.08	0.1	1.15	0.2	-0.07	1111011111
1559	9	10	2.26	1.07	1.11	0.1	1.40	0.4	-0.26	1111111110
1567	9	10	2.26	1.07	1.02	0.0	0.88	-0.1	0.19	1011111111
2795	9	10	2.26	1.07	1.09	0.1	1.26	0.3	-0.16	1111110111
2796	9	10	2.26	1.07	0.74	-0.3	0.41	-0.8	0.93	0111111111
2797	9	10	2.26	1.07	0.74	-0.3	0.41	-0.8	0.93	0111111111
2798	9	10	2.26	1.07	1.07	0.1	1.13	0.1	-0.05	1110111111
2799	9	10	2.26	1.07	0.74	-0.3	0.41	-0.8	0.93	0111111111
2800	9	10	2.26	1.07	0.74	-0.3	0.41	-0.8	0.93	0111111111
2801	9	10	2.26	1.07	1.08	0.1	1.15	0.2	-0.07	1111011111
2802	9	10	2.26	1.07	0.74	-0.3	0.41	-0.8	0.93	0111111111
2804	9	10	2.26	1.07	1.07	0.1	1.13	0.1	-0.05	1110111111
2805	9	10	2.26	1.07	1.08	0.1	1.15	0.2	-0.07	1111011111
2810	9	10	2.26	1.07	0.74	-0.3	0.41	-0.8	0.93	0111111111
2813	9	10	2.26	1.07	0.74	-0.3	0.41	-0.8	0.93	0111111111
983	1	10	-2.25	1.06	0.99	0.0	0.89	-0.1	0.16	0000001000
984	1	10	-2.25	1.06	1.02	0.0	1.01	0.0	0.04	0010000000
985	1	10	-2.25	1.06	0.97	0.0	0.83	-0.2	0.22	0000000100

2227	1	10	-2.25	1.06	1.01	0.0	0.97	0.0	0.07	0000100000
2231	1	10	-2.25	1.06	0.96	0.0	0.80	-0.2	0.26	0000000001
2232	1	10	-2.25	1.06	1.01	0.0	0.97	0.0	0.07	0000100000
2234	1	10	-2.25	1.06	0.96	0.0	0.80	-0.2	0.26	0000000001
2236	1	10	-2.25	1.06	0.96	0.0	0.80	-0.2	0.26	0000000001
2238	1	10	-2.25	1.06	0.99	0.0	0.89	-0.1	0.16	0000001000
2239	1	10	-2.25	1.06	0.96	0.0	0.80	-0.2	0.26	0000000001
3765	1	10	-2.25	1.06	0.96	0.0	0.80	-0.2	0.26	0000000001
3766	1	10	-2.25	1.06	1.06	0.1	1.27	0.3	-0.20	0100000000
3767	1	10	-2.25	1.06	1.01	0.0	0.99	0.0	0.05	0001000000
3768	1	10	-2.25	1.06	0.99	0.0	0.89	-0.1	0.16	0000001000
4724	1	10	-2.25	1.06	1.01	0.0	0.97	0.0	0.07	0000100000
4985	1	10	-2.25	1.06	1.06	0.1	1.27	0.3	-0.20	0100000000
2223	0	10	-3.53	1.85	MINIMUM ESTIMATED MEASURE					0000000000
2225	0	10	-3.53	1.85	MINIMUM ESTIMATED MEASURE					0000000000
4984	0	10	-3.53	1.85	MINIMUM ESTIMATED MEASURE					0000000000
MEAN	5.	10.	0.14	0.69	1.00	0.0	1.00	0.0		
S.D.	2.	0.	0.75	0.07	0.12	0.7	0.17	0.7		

ملحق رقم (7)

اختبار (t-test) لأداء الطلبة على الاختبارين السهل والصعب وفق نموذج راش

T-Test

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
IRTEASYO	5384	.8610	.9255	1.261E-02
IRT.HARD	5520	.1415	.7684	1.034E-02

One-Sample Test

	Test Value = 0					
					95% Confidence Interval of the Difference	
	t	df	Sig. (2-tailed)	Mean Difference	Lower	Upper
IRTEASYO	68.262	5383	.000	.8610	.8363	.8857
IRT.HARD	13.679	5519	.000	.1415	.1212	.1618

ملحق رقم (8)

معاملات الصعوبة لفقرات الاختبار من خلال أداء العينتين وفق نموذج راش

باستخدام برنامج (winsteps)

INPUT: 1490 PERSONS, 20 ITEMS ANALYZED: 1490 PERSONS, 20 ITEMS, 2 CATS

ITEMS STATISTICS: MEASURE ORDER

عينة الأداء المرتفع					عينة الأداء المنخفض				
ENTRY NUMBER	RAW SCORE	COUNT	MEASURE	ERROR	ENTRY NUMBER	RAW SCORE	COUNT	MEASURE	ERROR
9	619	1490	1.38	0.05	9	311	1490	1.25	0.06
5	968	1490	0.40	0.05	15	511	1490	0.56	0.05
15	970	1490	0.40	0.05	10	571	1490	0.38	0.05
2	1029	1490	0.21	0.06	17	601	1490	0.30	0.05
18	1040	1489	0.17	0.06	18	615	1490	0.26	0.05
17	1047	1490	0.15	0.06	6	626	1490	0.23	0.05
11	1069	1489	0.08	0.06	7	626	1490	0.23	0.05
10	1072	1490	0.07	0.06	5	635	1490	0.20	0.05
3	1088	1490	0.02	0.06	8	657	1490	0.14	0.05
6	1095	1490	-0.01	0.06	12	660	1490	0.13	0.05
1	1096	1490	-0.01	0.06	1	694	1490	0.04	0.05
16	1110	1490	-0.06	0.06	2	739	1490	-0.08	0.05
20	1112	1490	-0.07	0.06	3	759	1490	-0.14	0.05
12	1130	1490	-0.13	0.06	20	787	1490	-0.21	0.05
8	1134	1490	-0.15	0.06	11	807	1490	-0.27	0.05
7	1146	1490	-0.19	0.06	16	834	1490	-0.34	0.05
19	1161	1490	-0.25	0.06	19	848	1490	-0.38	0.05
13	1223	1490	-0.52	0.07	14	877	1490	-0.46	0.05
4	1263	1490	-0.71	0.07	13	1012	1490	-0.86	0.06
14	1274	1490	-0.77	0.07	4	1058	1490	-1.00	0.06

ملحق رقم (9)

اختبار(t-test) لأداء العينتين الفرعيتين المرتفع والمنخفض

وفق النظرية الكلاسيكية ونموذج راش

T-Test

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	HI.CTT	.7120	20	9.998E-02	2.236E-02
	LO.CTT	.4795	20	.1302	2.910E-02
Pair 2	HI.RASCH	5.000E-04	20	.4499	.1006
	LO.RASCH	7.000E-03	20	.4982	.1114

Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	HI.CTT & LO.CTT	20	.433	.057
Pair 2	HI.RASCH & LO.RASCH	20	.868	.000

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	Upper			
Pair 1	HI.CTT - LO.CTT	.2325	.1252	99E-02	.1739	.2911	8.306	19	.000
Pair 2	HI.RASC - LO.RAS	50E-03	.2484	54E-02	-.1228	.1098	-.117	19	.908